

A pH-dependent cluster of charges in a conserved cryptic pocket on flaviviral envelopes

Lorena Zuzic^{1,2}, Jan K Marzinek¹, Ganesh S Anand^{3,4}, Jim Warwicker⁵, Peter J Bond^{1,3*}

¹Bioinformatics Institute (A*STAR), Singapore, Singapore; ²Department of Chemistry, Manchester Institute of Biotechnology, The University of Manchester, Manchester, United Kingdom; ³Department of Biological Sciences, 16 Science Drive 4, National University of Singapore, Singapore, Singapore; ⁴Department of Chemistry, The Pennsylvania State University, University Park, United States; ⁵School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Institute of Biotechnology, The University of Manchester, Manchester, United Kingdom

Abstract Flaviviruses are enveloped viruses which include human pathogens that are predominantly transmitted by mosquitoes and ticks. Some, such as dengue virus, exhibit the phenomenon of antibody-dependent enhancement (ADE) of disease, making vaccine-based routes of fighting infections problematic. The pH-dependent conformational change of the envelope (E) protein required for fusion between the viral and endosomal membranes is an attractive point of inhibition by antivirals as it has the potential to diminish the effects of ADE. We examined six flaviviruses by employing large-scale molecular dynamics (MD) simulations of raft systems that represent a substantial portion of the flaviviral envelope. We utilised a benzene-mapping approach that led to a discovery of shared hotspots and conserved cryptic sites. A cryptic pocket previously shown to bind a detergent molecule exhibited strain-specific characteristics. An alternative conserved cryptic site at the E protein domain interfaces showed a consistent dynamic behaviour across flaviviruses and contained a conserved cluster of ionisable residues. Constant-pH simulations revealed cluster and domain-interface disruption under low pH conditions. Based on this, we propose a cluster-dependent mechanism that addresses inconsistencies in the histidine-switch hypothesis and highlights the role of cluster protonation in orchestrating the domain dissociation pivotal for the formation of the fusogenic trimer.

*For correspondence:
peterjb@bii.a-star.edu.sg

Competing interest: The authors declare that no competing interests exist.

Funding: See page 19

Preprinted: 13 July 2022
Received: 04 August 2022
Accepted: 18 April 2023
Published: 05 May 2023

Reviewing Editor: Qiang Cui, Boston University, United States

© Copyright Zuzic *et al.* This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editor's evaluation

Using state-of-the-art molecular dynamics simulations, the authors discuss the potential binding sites of drug molecules to the flaviviral envelope. Moreover, using constant pH simulations, they discuss the functional relevance of a cluster of ionizable residues in a cryptic site at the domain interface. These results have provided novel mechanistic insights into the pH-dependent conformational changes of the envelope protein and cryptic binding sites in the envelope protein that can be targeted for inhibiting viral infection.

Introduction

Flaviviruses belong to a family of enveloped positive-sense, single-stranded RNA viruses which are transmitted by arthropod vectors, predominantly mosquitoes and ticks. The family includes human

pathogens such as dengue (DENV), yellow fever (YFV), Zika, tick-borne encephalitis (TBEV), Japanese encephalitis (JEV) and West Nile virus (WNV), which can be subdivided into a neurotropic and non-neurotropic group of infectious agents (*Gaunt et al., 2001*). The non-neurotropic group of flaviviruses, associated with haemorrhagic disease, is primarily transmitted via *Aedes* mosquitoes and is mainly present in the tropical and subtropical regions of the world (*Bhatt et al., 2013; Kraemer et al., 2015*). In recent years, partly as a result of global warming, the mosquito distribution has been rapidly expanding into temperate regions and is now co-distributed with the majority of the human population (*Brady et al., 2012; Kamal et al., 2018; Kraemer et al., 2019*).

Flaviviral particles consist of an internal nucleocapsid and a surface-facing envelope, which is composed of a membrane and two types of embedded viral proteins. An envelope (E) protein is exposed to the surface of the virus and appears as a first point of contact between the virus and host immune system, whereas a smaller membrane (M) protein, which plays a role in viral maturation, localizes underneath the E protein and is concealed from the exterior. Both proteins are embedded in the viral membrane via transmembrane (TM) helices, and together they form stable heterotetrameric (E_2M_2) complexes (*Zhang et al., 2013a; Kostyuchenko et al., 2013; Sirohi et al., 2016; Sevvana et al., 2018*). In the mature virion, 90 E_2M_2 units are organised in a tight-knit herringbone-like pattern with three heterotetramers arranged in parallel to form a raft. In total, 30 rafts form a complete viral envelope and are likely a functional unit for immune recognition (*Zhang et al., 2013a; Sevvana et al., 2018; Figure 1*).

During the viral entry process, the virion is internalised via clathrin-mediated endocytosis (*Mosso et al., 2008; Acosta et al., 2008*) and is contained within endosomes. A gradually acidifying environment inside the compartment (\sim pH 5–6.5; *Randolph and Stollar, 1990*) triggers a large-scale conformational change within the envelope whereupon E proteins dissociate from their E_2 dimeric state, which lies flat on the surface of the virus, and instead form trimeric spikes with conserved fusion loops positioned at the tips (*Modis et al., 2004*). The fusion loops anchor the virus to the host membrane, effectively creating a bridge that spans the gap between the viral and endosomal membranes. A series of putative conformational changes in the spike trimer catalyse membrane fusion, penetrating the endosome and allowing for the viral nucleocapsid with its RNA genome to be released into the host cytoplasm (*Harrison, 2015*).

The identity of the exact residues responsible for triggering this dramatic sequence of conformational changes remains unclear, but the majority of research has been focussed on several conserved histidine residues as the most likely pH-sensing candidates, guided by the principles of the ‘histidine switch hypothesis’. This hypothesis presupposes histidines to be pH-sensing residues responsible for triggering conformational changes in viral fusion proteins that are dependent on pH. The hypothesis hinges on the property of the histidine side chain having its pK_a close to physiological pH ($pK_a = 6.0$), and therefore presumes that only histidines are able to change protonation state under the relatively mild pH changes occurring inside the living cell. Site-directed mutagenesis experiments on recombinant subviral particles of TBEV demonstrated that a single His317 mutation (residue numbers corresponding to DENV2) was sufficient for inhibiting a membrane fusion step, likely because of impaired dissociation of E protein dimers. Double His244/His282 mutants had a reduced stability of fusogenic trimers, which also impaired fusion (*Fritz et al., 2008; Stiasny et al., 2011*). A mutation in His144 could not yield stable subviral particles and its effect on the fusion step could therefore not be determined. However, subsequent research on WNV reporter virus particles has brought into question the histidine-switch hypothesis of the pH-dependent conformational change (*Nelson et al., 2009*). Even though mutations of His144 and His244 significantly affected the fusion process, substitutions to non-titratable Met/Asn and aromatic residues, respectively, resulted in partial rescue of infectivity.

One of the best researched flaviviruses is DENV (serotypes 1–4) which poses a serious health risk because it can lead to dengue haemorrhagic fever or dengue shock syndrome (*Gould, 1998*). More serious forms of the disease have been linked to antibody-dependent enhancement (ADE), where a first infection produces anti-DENV antibodies specific against one strain, which then imperfectly cross-react with another strain during a subsequent infection (*Halstead, 1979; Balsitis et al., 2010; Guzman et al., 2013*). This unexpected detrimental effect of the host’s immune system has been linked to subtle differences in E protein sequences among DENV serotypes, inefficient viral maturation resulting in a high number of immature particles presenting in the extracellular space, as well as to a striking heterogeneity of DENV particle morphology observed under various environmental

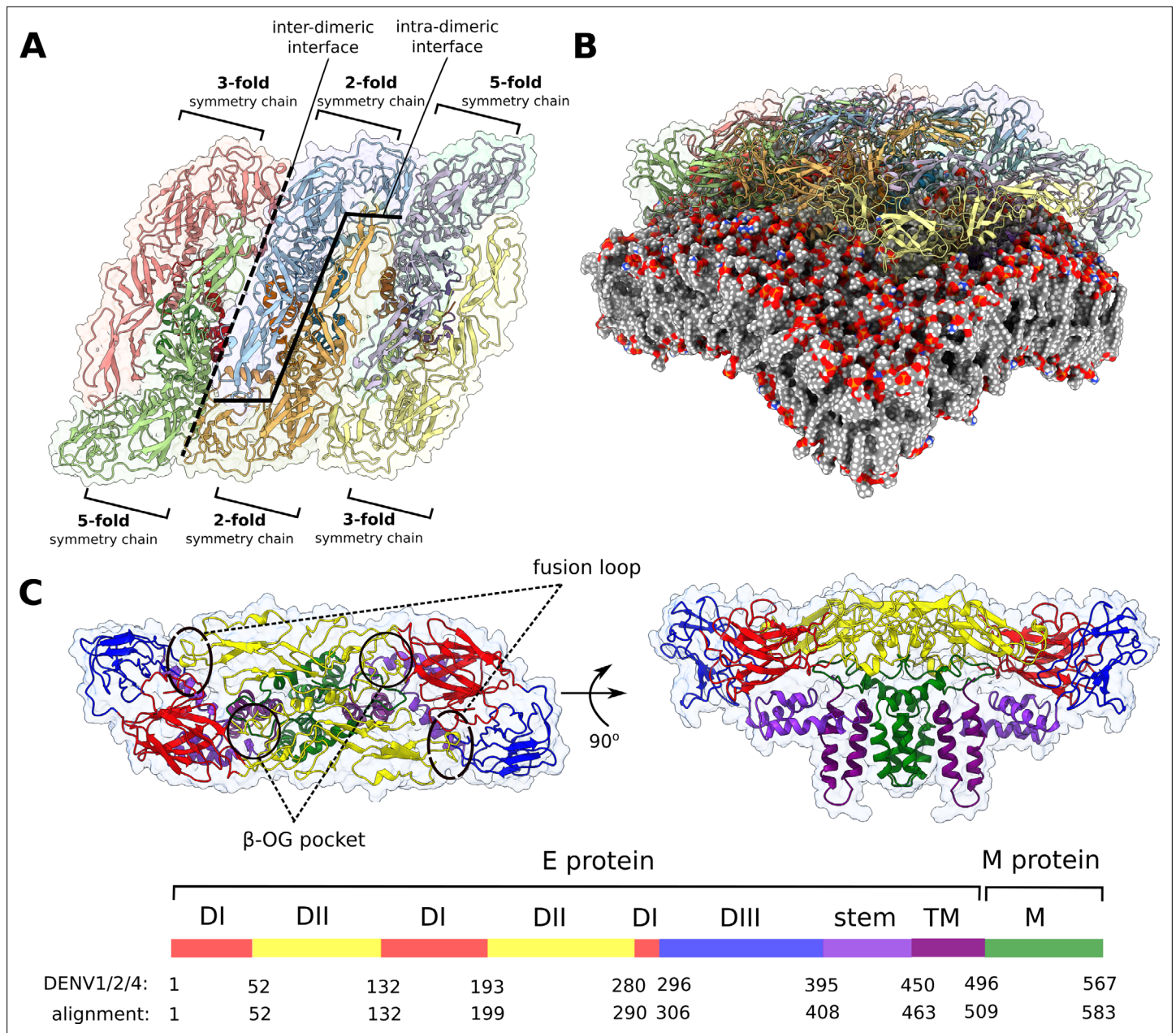


Figure 1. Flaviviral raft structure and domain organisation. **(A)** A top-down view on the flaviviral envelope protein raft consisting of three E_2M_2 heterotetramers organised in a parallel fashion. Chains differ in the context of their immediate environment and are described in terms of the symmetry fold, which is repeated across the entire viral particle. In this model, the majority of interchain contacts is retained for the central dimer (consisting of two 2-fold symmetry chains), while the edge 3-fold symmetry chains are not bordering neighbouring rafts as in the full virion. **(B)** The curvature of the envelope membrane is imposed by that of the protein rafts. See also **Figure 1—figure supplement 1**. **(C)** Top and side views of the E_2M_2 heterotetramer shown in ribbon representation and coloured by functional domains. Protein surface is shown as a transparent outline. The six flaviviral strains differ in length, and residue numbers on the key correspond to the alignment of all strains and to DENV1, 2, and 4. By convention, individual residues mentioned in the text are always matched with DENV2 numbering.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. An atomistic model of a DENV2 raft and its curved membrane overlaid with a cryo-EM map of a DENV2 envelope (PDB: 3J27; EMD: EMD-5520).

conditions associated with temperature and pH (Modis et al., 2004; Junjhon et al., 2010; Fibriansah et al., 2013; Zhang et al., 2013b; Lim et al., 2019; Morrone et al., 2020; Fibriansah et al., 2021). The mechanism of ADE is most likely multifaceted (Morrone and Lok, 2019; Wirawan et al., 2019), but predominantly includes anti-DENV antibodies interacting with the Fc γ receptor (Fc γ R) leading to enhanced viral internalisation into the endosome, from which the virus can escape by utilising its fusogenic spikes (Balsitis et al., 2010; Rey et al., 2018; Williams et al., 2013). The aggravation of disease caused by antibodies puts into question conventional routes of fighting viral infection, which usually involve vaccines designed to promote humoral immunity and a strong antibody response to the presence of a virus. Sanofi's *Dengvaxia*, the only dengue vaccine currently in use, has been shown to be serotype-dependent and can lead to ADE in those who have not contracted the virus prior to vaccination (Guy et al., 2017).

An alternative to antibody-reliant approaches for fighting viral infections are antiviral drugs which may act upon various stages of the viral life cycle. There have been numerous attempts to apply drug design to target flaviviruses, and they have usually focussed on the non-structural viral proteins involved in viral replication or polyprotein processing, as well as some host targets involved in viral assembly (reviewed in Boldescu et al., 2017). Recent studies have demonstrated that cyanohydrazone molecules bind to a pocket on the E protein to inhibit viral entry at micromolar concentrations for DENV2, Zika, and JEV (de Wispelaere et al., 2018; Li et al., 2019). The pocket in question is located at the domain I-domain II interface of the E protein, is cryptic in nature, and was fortuitously observed in crystal structures of the DENV2 soluble portion of the E protein (sE) bound to a single *n*-Octyl- β -D-Glucoside (β -OG) detergent molecule (Modis et al., 2003). The effectiveness of these antivirals in blocking the early stages of the viral life cycle without causing ADE encourages exploration of other possible druggable sites on flaviviral envelopes. It also highlights the potential of utilizing drug design approaches targeting cryptic sites that may be of functional importance to the virus.

One of the methods focussed on detection of potential drug-binding sites is a co-solvent simulation approach that exploits the property of preferential binding of small organic molecules onto protein surfaces with which they can establish favourable interactions. Such surfaces are known as hotspots and their detection is a critical step for rational drug design (DeLano, 2002). To that aim, the probes resembling common functional drug moieties (such as benzene or hydrocarbons) are preferably used, as their binding location and mode of interaction may be instructive for a subsequent drug design step (Guvench and MacKerell, 2009; Tan et al., 2016). On top of being a frequent component of drug-like molecules (Kolb and Caflich, 2006), benzene is also apolar in nature. This property is useful for uncovering cryptic sites which are at least partially hydrophobic in character and therefore only transiently appear in a detectable 'open' conformation (Cheng et al., 2007; Schmidtke and Barril, 2010; Oleinikovas et al., 2016; Kuzmanic et al., 2020).

Here, we use large-scale co-solute molecular dynamics (MD) simulations to explore the behaviour of six flaviviral rafts (DENV1-4, YFV, and Zika) and reveal conserved cryptic pockets in the presence of benzene probes. To that aim, we establish a raft model as a realistic representation of the viral envelope that successfully reproduces the features of curvature and interchain contacts, while being attainable in an all-atom level description and thus providing a precise insight into physicochemical properties applicable to the envelope as a whole. The dynamics of the ~400,000 atom raft systems are explored using a total simulation sampling time of over 14 μ s. We investigate a landscape of potentially druggable sites on the envelope surface, compare pocket conservation across viral strains in relation to their sequence similarity, and propose conserved and functionally important sites with potential for drug targeting and inhibition of the early stages of the viral life cycle. Finally, we elaborate on the functional importance of a proposed cryptic site by highlighting a conserved charged cluster of residues that includes the key histidine residue, His144, explore its behaviour in different pH-environments using a constant pH-simulation approach, and propose a cluster-dependent mechanism of DI-DIII dissociation essential for formation of the fusogenic trimer.

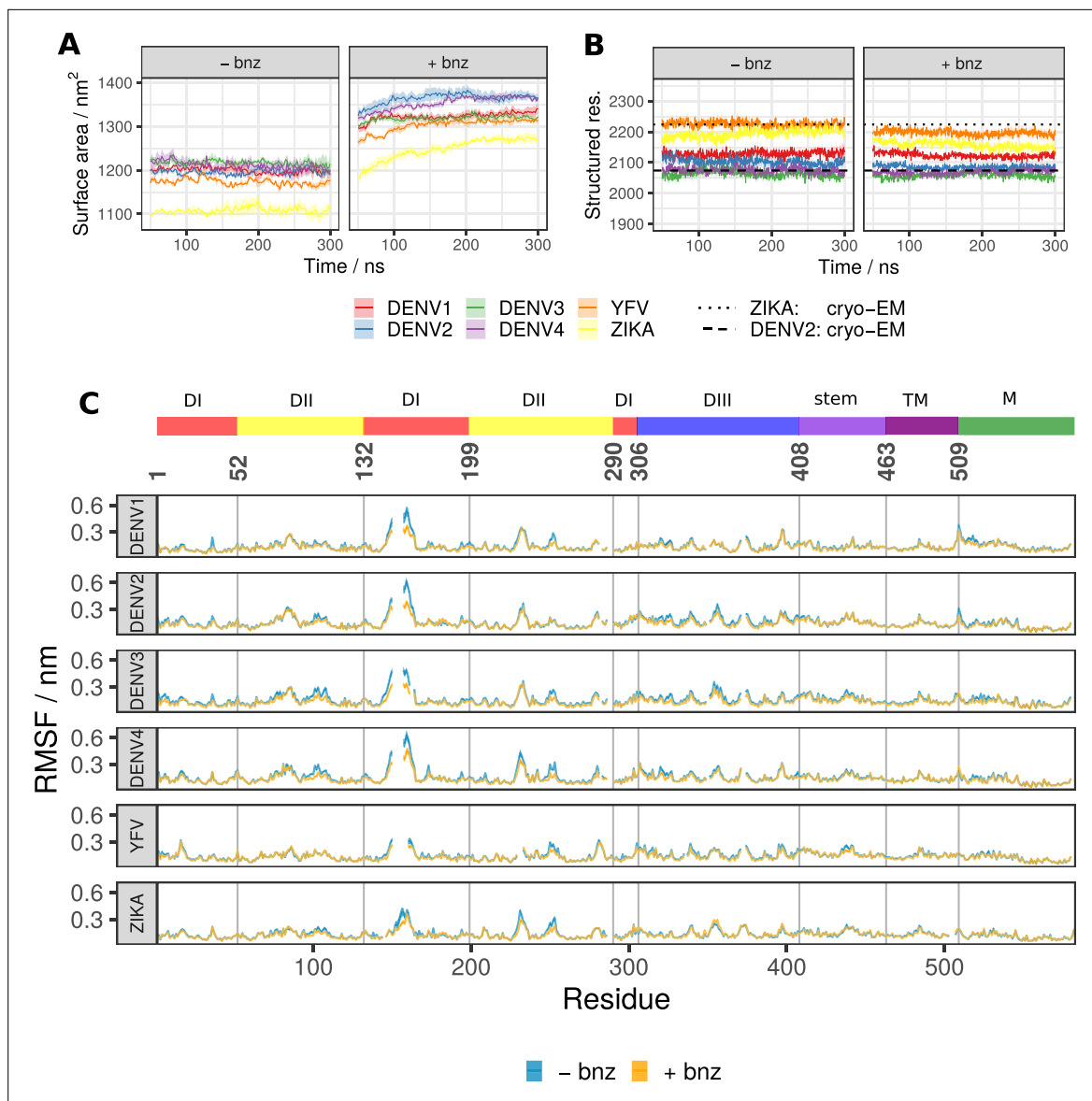


Figure 2. Global raft properties for all six viral serotypes showing the effect of benzene probes added to the solvent. **(A)** The total SASA across all viral systems increases with the addition of hydrophobic benzene probes to the solvent. The first 50 ns of each simulation was considered as equilibration and therefore omitted from the analysis. **(B)** The number of residues participating in secondary structure (defined as constituent elements of α -helices, β -sheets, β -bridges, or turns) remains largely unaffected in benzene systems and closely corresponds to the values describing the experimentally derived cryo-EM structures. **(C)** The RMSF values describing whole-residue fluctuations are similar across solvent types and viral serotypes. The gaps in lines are due to the sequence alignment which accounts for differences in length or deletions/insertions between different flavivirus envelope proteins, ensuring that the corresponding residues are vertically aligned. All values displayed in panels are averaged across repeats and, in the case of RMSF, across all chains of the system. Standard error is shown as a transparent ribbon around the mean value.

Results and discussion

Benzene binding increases solvent protein accessibility without affecting the secondary structure

High concentrations of apolar probes can potentially lead to undesirable unfolding events by invading and disrupting the hydrophobic core of the protein, resulting in a shift towards biologically irrelevant conformations (Schmidt *et al.*, 2019). Thus, in the current work, we first confirmed that the addition of benzene to flaviviral raft systems increases the total SASA without affecting protein secondary structures (Figure 2A–B). The unperturbed secondary structure suggests that the increased SASA

of the sampled conformations of the rafts is not due to protein unfolding, but instead is a result of subtle side chain rearrangements, benzene incorporation into protein-protein interfaces, and cryptic pocket opening events. Additionally, very similar RMSF profiles for water-only and benzene simulations demonstrated that the overall dynamic fluctuations were comparable in behaviour despite the differences in solvent composition (**Figure 2C**). Highly dynamic loop regions (such as the N153 glycan loop) had fractionally lower RMSFs compared to water-only simulations. This effect might be due to stabilisation imposed by benzene interacting with the disordered loop regions and the ordered surface beneath, thus providing a stable interaction surface for the flexible portions of the protein.

Benzene binding patterns are sequence-dependent

Flaviviral serotypes, although similar in structure, exhibit subtle serotype-specific differences in envelope dynamics. Even though the polyprotein sequence similarity between the presently explored flaviviruses is relatively high (42% for all six serotypes and 71% for DENV serotypes only), a large number of mutations localise on the surface-exposed portions of the envelope protein (**Figure 3—figure supplement 1A–B**). This is unsurprising in the context of viral evolution, as mutations arise in the surface-exposed areas in order for the virus to evade the host immune response. However, certain exposed regions of the E protein are highly conserved because they play specific functional roles in the viral life cycle and are therefore less likely to accumulate mutations.

Shared sites of benzene binding reflect residue conservation patterns on the E protein and mostly localise around the conserved fusion loop and at the pr-binding interface (**Figure 3** and **Figure 3—figure supplement 1C**). The fusion loop (residues 98–111) is conserved among flaviviruses as it inserts into the endosomal membrane, which is essential for catalysing the membrane fusion event leading to genome release into the host cell cytoplasm (**Allison et al., 2001**). Similarly, the pr portion of the prM protein protects the fusion loop from premature insertion during the maturation process, and its interaction surface on the E protein is therefore also conserved to a high degree (**Heinz et al., 1994; Figure 3—figure supplement 1A and C**). It is interesting to note that protein-protein interfaces have similar characteristics to drug-binding hotspots (**DeLano, 2002; Mattos and Ringe, 1996; Zerbet et al., 2012**) and the methods used for detecting one are also likely to detect the other type of protein surface. This was well demonstrated by benzene mapping, as we found that benzene molecules preferentially interacted with the pr-binding interface, highlighting it as a potential hotspot.

Overall, though, the surface patterns of benzene binding and, consequently, increased SASA, were variable and serotype-dependent (**Figure 3**). The external E protein surface is mostly unconserved and correspondingly showed a significant degree of variability in benzene binding patterns across serotypes. Notably, a central portion of DII was benzene-depleted (conservatively, residues 58–65, 120–125, 215–235, 252–267), suggesting that it is unlikely to be an attractive drug-binding hotspot. The width of this depleted band was sequence-dependent, with it being the narrowest in DENV4 and the widest in YFV. There are currently no known flaviviral antibodies that specifically target the central band of the raft, corroborating the observation that poor benzene binding sites are also unlikely to be involved in the formation of protein-protein interfaces. The benzene interaction sites slightly differed depending upon the position of the chain within the raft (**Figure 1A**). Protein-protein interfaces that were positioned on the raft edges were more easily accessible to benzene molecules than the interfaces engaged in interchain interactions. Despite this, benzene mapping was highly similar across chains and displayed clear serotype-specific patterns of binding.

The areas of frequent benzene binding also matched increases in SASA, showcasing the role of benzene in exposing protein sites that predominantly remain hidden in water-only simulations. The TM regions of the E proteins and the M proteins showed little increase in SASA, which is explained by the fact that those regions are sequestered within the membrane and are, for the most part, out of reach of benzene molecules.

Correlation analyses of serotype properties, especially SASA, recapitulated phylogenetic relationships between the viruses (**Figure 3—figure supplement 1D**). Serotypes that are more closely related generally exhibited more similar dynamic behaviour, highlighting the importance of sequence conservation in yielding shared characteristics such as SASA, benzene binding, or pocket formation. Interestingly, not all residue-based properties showed an equally high degree of correlation with sequence – notably, the RMSF correlations were poorly reflective of flaviviral evolutionary relationships, suggesting a more complex interplay of properties causing residue fluctuations than sequence

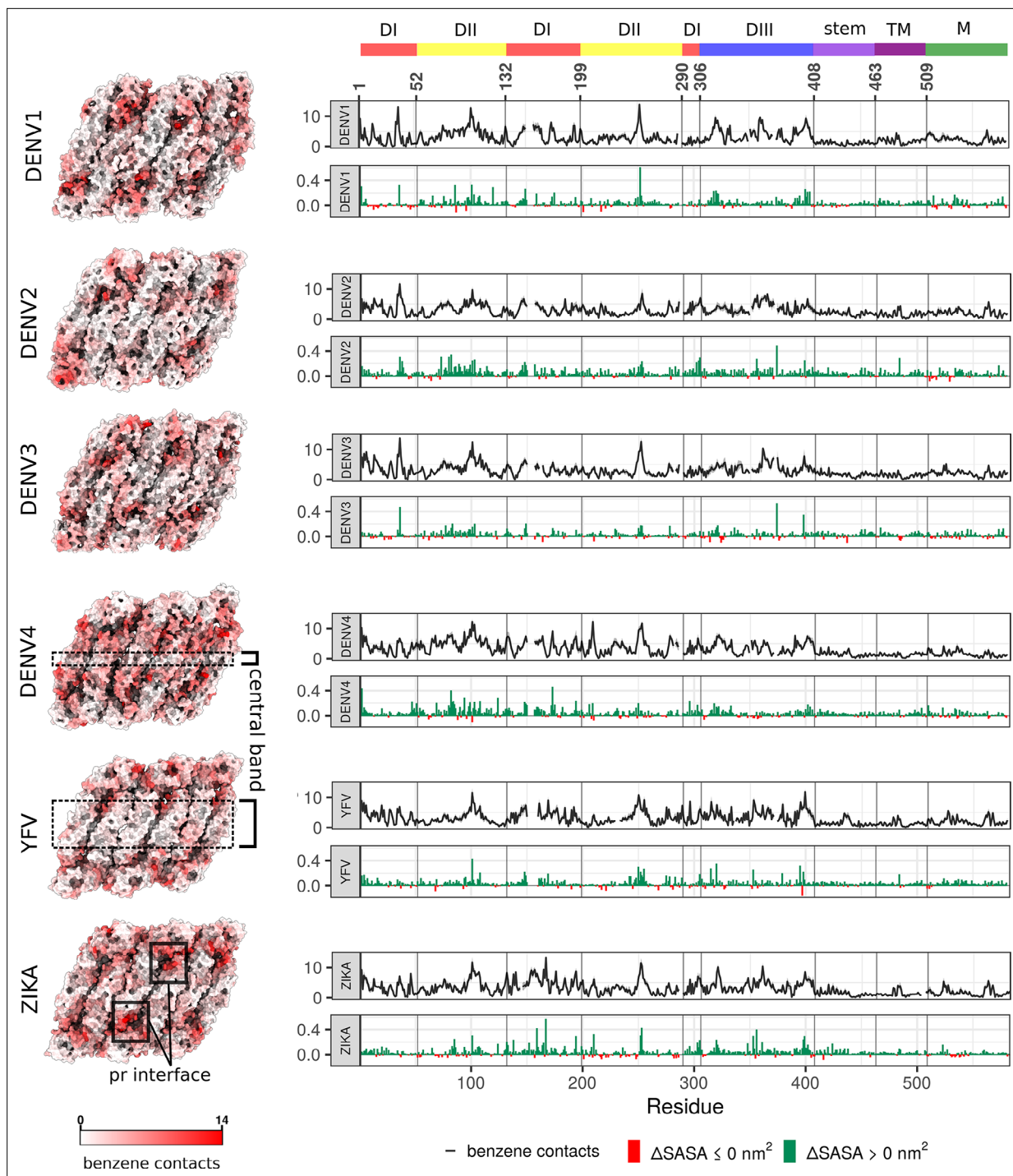


Figure 3. Benzene contacts and changes in SASA shown for all six viral serotypes. **(Left)** The average number of benzene contacts per residue is shown on viral raft structures. The values for individual chains are averaged across repeats. For patterns of residue conservation and phylogenetic relationships between strains, see **Figure 3—figure supplement 1**. **(Right)** In each upper graph, the mean value of benzene contacts per residue averaged across all chains and repeats is shown in black line, with standard error displayed as a grey ribbon. In each lower graph, the average change in SASA is shown in nm (Bhatt et al., 2013), calculated as $(SASA_{+bnz} - SASA_{-bnz})$, with an increase shown in green and decrease in red. As with **Figure 2**, the gaps in lines are due to flavivirus-specific sequence alignment.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Residue conservation and phylogenetic relationships of flaviviruses correlated with simulation properties.

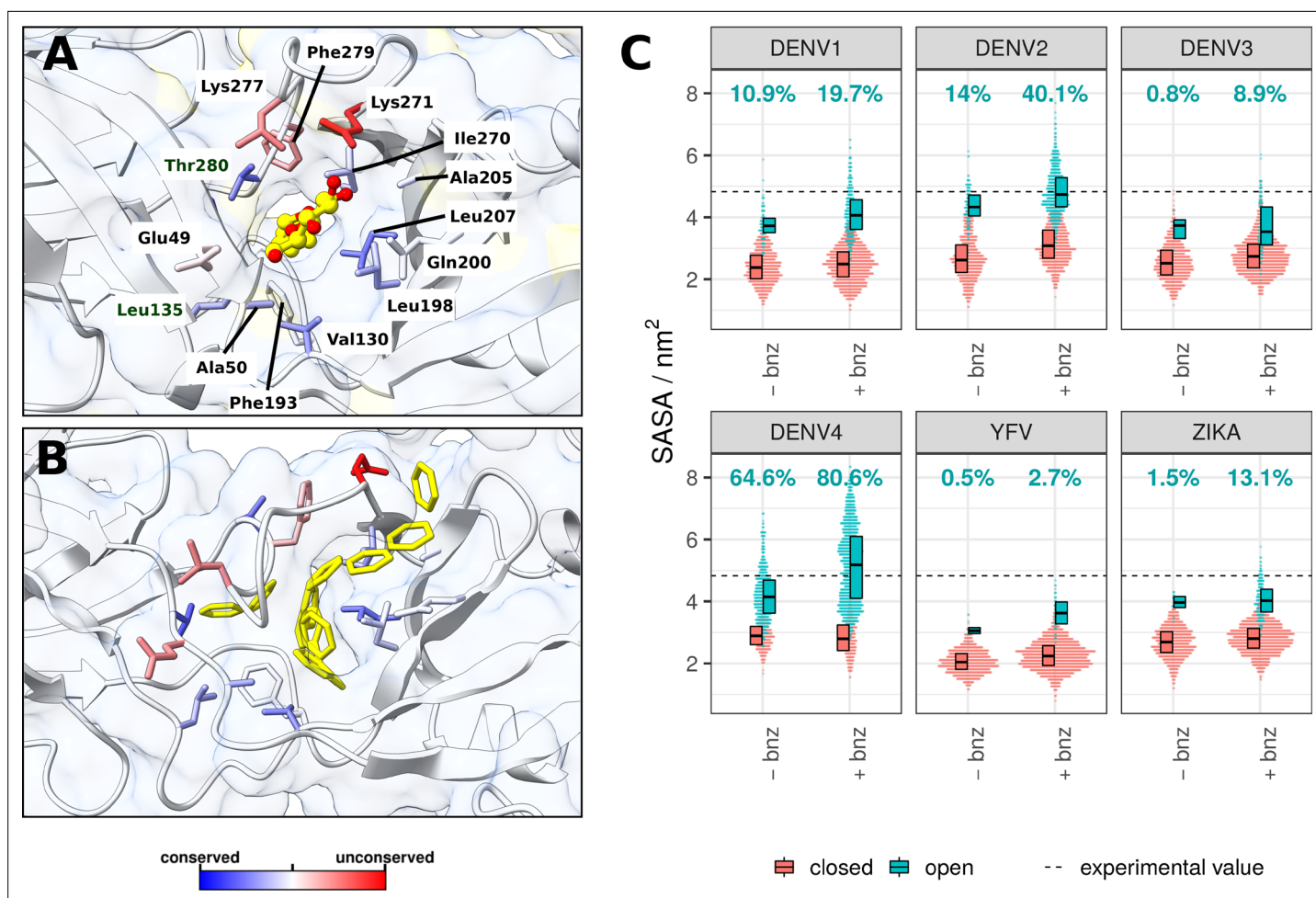


Figure 4. β -OG pocket in the presence of the β -OG ligand or benzene, and its SASA across viral serotypes. **(A)** The β -OG detergent molecule co-crystallised inside the cryptic pocket of the DENV2 sE protein (PDB code: 1OKE) (Modis et al., 2003). The protein is shown as a combination of white ribbon and transparent surface representations. The β -OG molecule is shown in yellow sticks, with oxygen atoms highlighted in red. The residues surrounding the ligand are accentuated in stick representation and coloured based on their conservation scores calculated in ConSurf (Ashkenazy et al., 2016). Residue labels correspond to DENV2. **(B)** A representative simulation snapshot of benzene molecules occupying the β -OG pocket in the DENV2 system. Benzenes are shown in yellow stick representation. Other elements of the system are visualised in the same way as in panel A. **(C)** Cumulative SASA for all residues lining the pocket (labelled in panel A) across six flaviviral serotypes. The data was visualised as a swarm plot, where a SASA measurement for each sampled frame was represented as a horizontally stacked non-overlapping point. The spread of points therefore describes the distribution of SASA measurements for systems in absence (-bnz) or presence of benzene (+bnz). Pocket SASA values are shown separately depending on the solvent composition. Open and closed conformations of the pockets were differentiated by applying a k -means clustering algorithm (details in the Methods section). The percentage of the β -OG pocket observed in an open conformation is specified for each simulation system. The indicated experimental SASA value of the pocket is that measured for the crystal structure (panel A) averaged across the two chains. See also Figure 4—figure supplement 1.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. β -OG pocket properties of SASA and benzene contacts shown for individual viral strains and chain types.

alone (Figure 3—figure supplement 1D). Instead, the RMSF values showed remarkable similarity across all six viral serotypes (Figure 2C).

The open/closed state ratio of β -OG pocket is specific to the viral serotype

The β -OG cryptic pocket was first detected in DENV2 sE protein crystal structures (Modis et al., 2003) where a single detergent molecule intercalated within the hydrophobic pocket located at the DI-DII interface of each chain (Figure 4A). The binding pocket was absent in apo-structures as the position of

the *kl* loop obscured the entry to the site (**Figure 1C**). Functionally, the DI-DII interface is a hinge region and its movement is essential for the E protein to undergo pH-dependent conformational changes leading to the formation of fusogenic trimers (**Zhang et al., 2004**). However, the co-crystallised β -OG molecule was absent from sE dimer structures of other flaviviruses (**Kanai et al., 2006; Nybakken et al., 2006**), and even from the E protein structures of the same viral serotype (**Zhang et al., 2004**) despite the addition of the detergent in the crystallisation buffer. Nonetheless, other studies have shown that the β -OG pocket is likely present in at least some flaviviruses. The JEV apo-sE crystal structure showed that the *kl* loop is positioned in a way to reveal a pocket large enough to accommodate β -OG, even in the absence of the ligand (**Luca et al., 2012**). Indeed, pyrimidine compounds and cyanohydrates (both groups exhibiting inhibitory effects on the fusion events of DENV2, WNV, JEV, and Zika viruses) appear to bind inside the β -OG pocket, as demonstrated by loss-of-binding site-directed mutagenesis and photocrosslinking experiments (**de Wispelaere et al., 2018; Li et al., 2019**).

Using the benzene mapping simulation approach on raft systems, we explored the characteristics of the β -OG pocket in different flaviviral serotypes. We observed that the ratios of open/closed states of the pocket differ greatly depending upon the particular virus (**Figure 4C**), with the biggest pocket exposure attributed to DENV4 followed by DENV2, with YFV exhibiting comparatively the lowest pocket SASA even in the presence of benzene. Pocket properties were also affected by solvent composition, with simulations containing benzene manifesting, on average, higher SASA values. Benzene molecules can enter the hydrophobic pocket and transiently interact with residues lining the site (**Figure 4B**). The number of benzenes that occupied the pocket during the course of the trajectory was in direct correlation with the pocket SASA (**Figure 4—figure supplement 1B**), with DENV2 and DENV4 showing the highest propensity for accommodating benzene molecules within the pocket, both in terms of number and frequency. In contrast, YFV β -OG pockets were the least likely to contain any benzene. Therefore, the presence of benzene probes in the system modified pocket exposure and, consequently, its SASA, either by locking the pocket in an open conformation or by inducing the opening of the pocket in the first place.

The observed open/closed ratios of the β -OG pocket were not equivalent for all chains in the system; instead, they were dependent on the position of the chain within the raft. The edge 3-fold chains appeared to have the highest median SASA value, an effect that was in particular exaggerated for DENV2 and DENV4 systems (**Figure 4—figure supplement 1A**). This might be due to the fact that the edge chains were unconstrained by the neighbouring elements on one side, allowing for greater flexibility of the protein and easier opening of the pocket.

We clustered pockets based on SASA and contact properties to separate the sampled conformations into two clusters corresponding to open and closed states (**Figure 4C**). The DENV4 pocket appeared to be predominantly sampled in an open conformation; the DENV2 pocket was in an open state only half as often as compared to DENV4. Contrastingly, DENV1 and DENV3 exhibited very few open conformations. This variability is reflective of a relatively low conservation of residues (53%) among the DENV group of viruses – an attribute that also highlights the potential difficulty in developing pan-flaviviral therapeutics targeting the β -OG site.

The evident differences in pocket exposure can only be partially explained by the differences in the residues lining the pocket, and it is likely that the pocket opening is influenced in the wider context of protein residues outside of the binding site. **de Wispelaere et al., 2018** identified a mutation in DENV2 that is detrimental to antiviral compound binding, even though the point-mutation M196V is outside of the β -OG pocket. These findings suggest that the pocket opening might be allosterically affected by residues outside the binding cavity, either by modifying its shape, frequency of opening, or by affecting the stability of the interactions with its binding ligands.

Remarkably, the SASA value measured for the open β -OG pocket in the crystal DENV2 sE structure (**Modis et al., 2003**) closely corresponded to the mean pocket value in simulated DENV2, suggesting that the benzene simulations successfully sampled biologically relevant open conformations of the pocket that match experiment. Despite some of the serotypes exhibiting predominantly closed states of the pocket even with the addition of benzene, those β -OG pockets are not necessarily absent. Instead, the kinetics of site opening might be serotype-dependent, resulting in pockets of some of the viruses predominantly residing in closed states. Thus, successful inhibition of the fusion events in Zika in the presence of β -OG pocket-binding inhibitors suggests that the pocket is sufficiently druggable, despite its presently observed low fraction (13.1%) of open states.

The α pocket located on the domain interfaces is conserved and contains a buried cluster of charges

Our previous studies of a single DENV2 E₂M₂ heterotetramer identified a novel cryptic pocket (referred to as α) located on domain interfaces (Zuzic *et al.*, 2020; Figure 5). Here, we found that this pocket was consistently present in all explored flaviviral serotypes and for all raft chains (Figure 5—figure supplement 1). The increase in SASA in the presence of benzene was evident for all mapped α pockets, confirming its cryptic character. Interestingly, properties of the α pocket across serotypes appeared to be more consistent than those of the β -OG pocket, and this uniformity could be linked to a higher degree of residue conservation (47% for all explored flaviviruses; 70% for the DENV group). SASA was also less affected by the location of the chain within the raft for the α pocket, likely due to the fact that it is positioned peripherally and is therefore less dependent on intra-raft contacts. We cannot exclude that the dynamics of the α pocket opening might be affected by inter-raft contacts not present in our models.

The pocket is elongated and functionally divided into two parts - the fusion loop interface, representing the interchain interaction surface; and the DI-DIII cryptic pocket segment that is located underneath the N153-glycosylation loop (Figure 5A). A closer inspection of the cryptic segment reveals that it is well conserved, with surface residues showing a greater degree of variability, and buried residues appearing overall more conserved. The N153-glycan loop obstructing the entrance to the pocket was highly flexible in our simulations for all DENV serotypes (Figure 1C). Although we omitted the glycan component from our models, the simulations were in agreement with the hydrogen-deuterium exchange experiments performed on the whole DENV2 particle that demonstrate a high degree of flexibility in the glycosylation loops (Lim *et al.*, 2017). Furthermore, cryo-EM structures of the whole viral particles (Zhang *et al.*, 2013a; Sevvana *et al.*, 2018; Renner *et al.*, 2021), as well as crystal structures of E proteins expressed in insect cells (Modis *et al.*, 2003; Zhang *et al.*, 2004; Modis *et al.*, 2005; Barba-Spaeth *et al.*, 2016) indicate that the N153-loop is unstructured. Interestingly, E protein dimers expressed in bacterial systems (Luca *et al.*, 2012; Lu *et al.*, 2019) appear to have the N153-loop structured as an α -helix, suggesting a role of the glycan moiety in loop secondary structure organisation. The flexibility of the N153-loop indicates that it might aid cryptic pocket opening and transient solvent accessibility of residues lining the α pocket.

The DI-DIII cryptic segment of the pocket contains a conserved His144 residue (Figure 5A) that has been implicated in fusogenic pH-dependence (Nelson *et al.*, 2009) and viral particle stability (Fritz *et al.*, 2008). In a static structure (Zhang *et al.*, 2013a), the residue appears to be solvent-inaccessible and with a predicted pK_a outside the biological range (pK_a < 0), which would make it an unlikely candidate for a pH-sensing residue (Figure 6—figure supplement 2A). However, site-directed mutagenesis experiments indicate its importance in the pH-dependent conformational change (Nelson *et al.*, 2009), although its exact role is as of yet unknown. The dynamic exploration of the flaviviral envelopes in our simulations shows that the residue could be transiently exposed to the solvent, concurrent with the cryptic pocket opening events, which in turn might affect its pK_a.

Generally, the only residues that have tended to be considered in the context of pH-sensing roles in the flaviviral E proteins have been histidines as their intrinsic sidechain pK_a values (~6.0, as determined from titration experiments of compounds outside the protein environment) can rapidly protonate under relatively mild acidic conditions. However, this approach disregards the effect of the surroundings on residue pK_a (García-Moreno *et al.*, 1997; Isom *et al.*, 2011) and the fact that the residues, especially buried ones, are located in a low dielectric environment. Under such conditions, pK_a values can be dramatically altered – specifically, Asp/Glu residues have the potential to be upshifted (Srivastava *et al.*, 2007) and act as pH-sensing or pH-coupled residues (Sazanavets and Warwicker, 2015). When E and M protein residues were explored in the context of their burial, conservation, and potential for being titrated, only six of them fulfilled the criteria and, remarkably, all but one were located on the DI-DIII interface associated with the α pocket region (Figure 5B). Four conserved residues – Arg9, Asp42, His144, and Glu368 – were found to closely interact and formed a distinct cluster of charges bridging the DI-DIII interface. In addition, the N-terminus (N-ter) – characterised by its burial and ionisability, and unaffected in terms of conservation – was in close proximity to the residues in question and was also a part of the cluster. The presence of closely coupled residues within the cluster suggests that any pH-dependent protonation event on His144 would have to be considered in the context of other charges

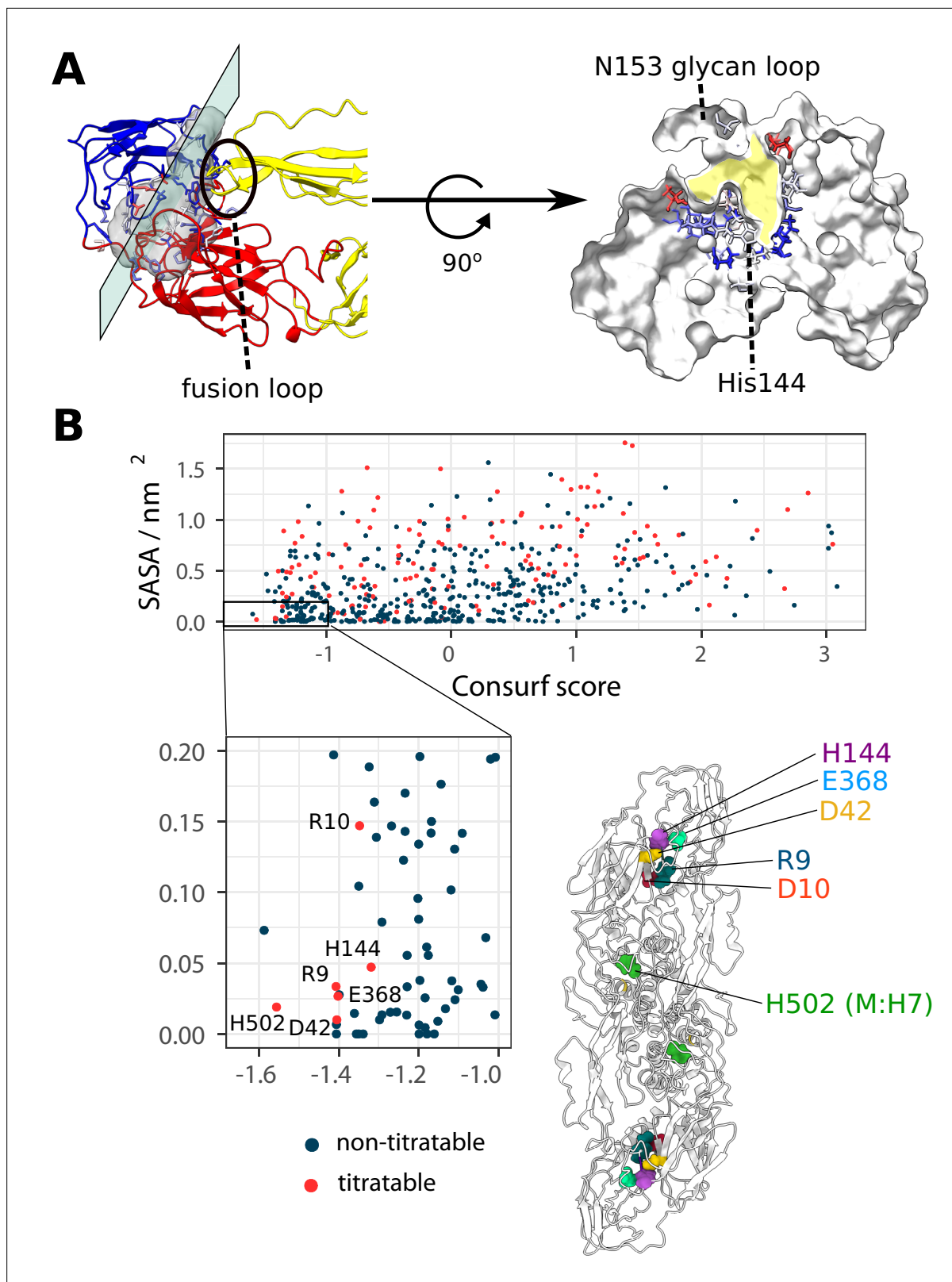


Figure 5. The location of the α pocket and the selection of all conserved, buried and titratable residues on the flaviviral envelope, showcasing a charge cluster associated with the pocket. (A) Left side: The α pocket density displayed as a transparent grey surface located on the intrachain DI-DIII and interchain DII-DIII interface. The E protein is shown in ribbon representation and coloured according to domain assignments (as seen in **Figure 1**). The plane represents a directional slice of the pocket shown on the right side of the panel. Right side: The protein is shown in white surface representation

Figure 5 continued on next page

Figure 5 continued

and the residues lining the pocket are displayed as sticks and coloured according to their ConSurf conservation score. The pocket is highlighted in yellow. (B) The selection of E and M protein residues according to the properties of conservation, burial, and ionizability. Residues characterised by a high degree of conservation (ConSurf score ≤ -1.0) and solvent inaccessibility (SASA ≤ 0.2 nm²) are shown in the inset. Additionally, titratable residues are labelled and displayed on the E₂M₂ structure, where it becomes apparent that almost all selected residues occupy the DI-DIII interface region associated with the α pocket. For α pocket SASA measurements across flaviviral strains, see **Figure 5—figure supplement 1**.

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. α pocket SASA values across different viral strains and solvent compositions.

within the cluster. The effects of pH on the cluster protonation and dynamic behaviour were next explored in constant-pH simulations.

Disruption of the α pocket charge cluster at low pH

We performed 0.88 μ s of constant-pH simulations on the sE dimer of DENV2 across a broad pH range of 0–9. The convergence of the examined pK_a values is shown in **Figure 6—figure supplement 1**. The focus of our analysis was the α pocket charged cluster (N-ter, Arg9, Asp42, His144, Glu368) and its behaviour at different pH conditions (**Figure 6**). Interestingly, even though preliminary pK_a calculations based on the static structure obtained from cryo-EM predicted His144 to be inaccessible for protonation (**Figure 6—figure supplement 2A**), conformational sampling during constant-pH simulations revealed transient solvent accessibility of the charge cluster that likely allows for histidine protonation at endosomal pH (pK_a(His144)=4.84; **Figure 6—figure supplement 2B**).

The disruption of cluster contacts (assessed by measuring the radius of gyration of the cluster residues, with higher values corresponding to overall greater disorder) was assessed in relation to the patterns of charges. Considering that the charge cluster is broken apart in the postfusion trimer (**Modis et al., 2004**), it is expected that a decrease in pH would be accompanied by disruption of the cluster. The cluster underwent a charge shift as consequence of lowering pH (**Figure 6B**), which indeed led to a gradual increase in disorder (**Figure 6A**), most noticeable following a total cluster charge shift from 0 to +3.

In high and neutral pH environments, the cluster was mainly in the D⁰H⁰E⁻ state, corresponding to protonated N-ter and Arg9 and deprotonated Asp42, His144 and Glu368, resulting in a total cluster charge of 0. The residues within the cluster were predominantly in close contact and formed a tightly interacting unit on the DI-DIII interface. When assessed across all pH values, this close-contact state dominated the conformational landscape as it occurred in 95% of all sampled structures (**Figure 6C**). Notably, the remaining conformations, characterised by looser residue association, were enriched at positive cluster charges occurring at low pH (**Figure 6—figure supplement 3**). This uneven distribution of conformations across charge states suggests a dissociation effect linked with acidic pH conditions, despite its relatively low occurrence as compared to the compact conformation. The constant-pH simulations did not include titrations of the N-ter or Arg9 and instead contained a fixed +1 charge at all pH values. However, given that the intrinsic pK_a values for both residues are in the basic range (pK_a = 9.6 and pK_a = 12.0) and their predicted pK_a values are 7.4 \pm 0.2 and 9.45 \pm 1.9, respectively (**Figure 6—figure supplement 2A**), it is likely that both residues remain largely protonated in the acidic environment of the endosome.

Specific patterns of charge distribution across residues had different effects on cluster stability, even when the total charge remained unchanged: D⁰H⁰E⁻ (protonated His144) had little effect on cluster disruption, while D⁰H⁰E⁰ (protonated Glu368) introduced a charge imbalance effect leading to the disturbance of the overall cluster structure. The cluster at +1 charge was a significant presence in the pH 5–6 range, relevant in the context of the endosomal pH-dependent conformational change (**Figure 6B**).

Transition of the total cluster charge to the +2 value resulted in substantial destabilisation of the DI-DIII interface, specifically in a D⁰H⁰E⁰ state (protonated His144 and Glu368). The switch to this charge state was accompanied by a protein-wide conformational change in the sE dimer with DIII shifting upwards from its initial position (**Figure 6C**). The disruption was linked to Arg9 decoupling from the rest of the residues to point away from the cluster. At very low pH (0–2), the cluster exhibited a total charge of +3 (D⁰H⁰E⁰, corresponding to all-residue protonation), and this was associated with even greater disruption of cluster contacts and conformational changes that manifested as DIII

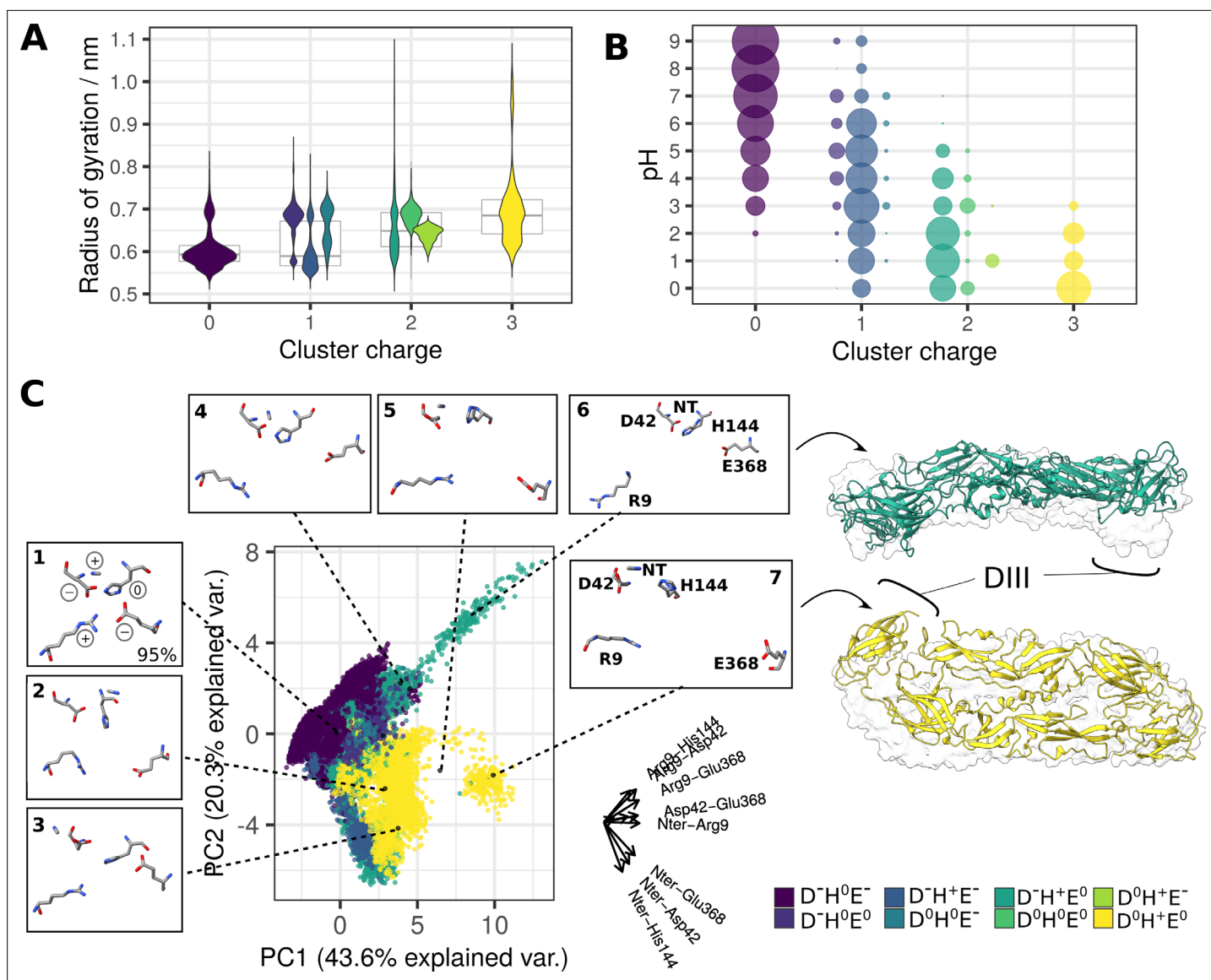


Figure 6. The response of the charge cluster to varying pH conditions. For convergence of pK_a values in constant-pH simulations, see **Figure 6—figure supplement 1**. **(A)** Disruption of the charge cluster (assessed via its radius of gyration) in relation to the charge pattern (violin plots in the foreground) and overall cluster charge (boxplots in the background). The measurements were obtained over a range of pH values (pH 0–9). The violin plots represent the distribution of y -values (radius of gyration measurements) for each protonation state, where the total area of violin(s) for each protonation state is constant. The codes in the legend describe residue charges for the residues that were titrated in the simulation system: Asp42, His144, and Glu368. N-ter and Arg9 had a fixed positive charge and were therefore omitted. For predicted pK_a values of selected residues, see **Figure 6—figure supplement 2**. **(B)** Occurrence of distinct protonation states at different pH conditions. Point sizes correspond to the population sizes at each value. $D^-H^0E^-$ is a dominant pattern of charge in high and neutral pH conditions, whereas the remaining seven patterns (representing +1,+2, and +3 total cluster charge) appear at low pH conditions. See also **Figure 6—figure supplement 3**. **(C)** PCA describing cluster conformations based on the intra-cluster distances. The axes corresponding to the original variables are shown on the right side of the plot (labels for the smallest axes His144-Glu368 and Asp42-His144 omitted for clarity). Clustering analysis of the five residues (details in the Methods section) generated nine distinct conformations which are projected onto the PCA and shown in insets around the plot. Conformations 8 and 9 had a negligible contribution to the total number of structures and were therefore excluded from the visualisation. The snapshots corresponding to structures undergoing a greater conformational change at low pH are visualised in ribbon representation (trajectories shown in **Figure 6—videos 1; 2**). Transparent surfaces represent the starting sE dimer conformation, demonstrating the extent of the conformational shift. Ribbon colours correspond to the specific charge patterns under which the conformational change had occurred (green and yellow for $D^-H^+E^0$ and $D^0H^+E^0$, respectively). For SASA of the α pocket across different charge states of the cluster, see **Figure 6—figure supplement 4**. For the effects of benzene on the conserved cluster behaviour, see **Figure 6—figure supplement 5**.

The online version of this article includes the following video and figure supplement(s) for figure 6:

Figure supplement 1. Convergence of pK_a values.

Figure 6 continued on next page

Figure 6 continued

Figure supplement 2. Prediction of pK_a values for selected E protein residues using static structures and MD simulations.

Figure supplement 3. Occurrence of conformations at different charge states.

Figure supplement 4. α pocket SASA values across different protonation states (violins) and charge states (boxplots) of the conserved cluster.

Figure supplement 5. Benzene contacts and its effects on the charge cluster.

Figure 6—video 1. The top-down view of the DIII dissociation in DENV2 sE dimer under low pH conditions.

<https://elifesciences.org/articles/82447/figures#fig6video1>

Figure 6—video 2. The side view of the DIII tilt in DENV2 sE dimer under low pH conditions.

<https://elifesciences.org/articles/82447/figures#fig6video2>

shifting away from DI (**Figure 6C**). Notably, this change appears to be different from the conformational switch observed in the $D^+H^0E^0$ state as it was predominantly linked to Glu368 distancing from the remainder of the cluster residues.

Even though we observed protein-wide conformational changes only in more extreme pH environments than are observed within the endosome, it is evident that at least one associated pattern of charge ($D^+H^0E^0$) can occur at pH 5, suggesting that this route of DI-DIII disruption at endosomal pH is possible. Alternatively, the observed increase in intra-cluster distances at total charge zero suggests that it might also lead to DI-DIII dissociation in the context of better sampling or a more expansive representation of the viral particle surface.

Based on our findings, we propose that the drivers of conformational change at low pH are more complex than the simplistic histidine switch hypothesis, which limits possibilities for pH-sensing residues exclusively to histidines. As shown in our simulations, the solitary His144 protonation event has little effect upon overall sE dimer stability; instead, the conformational change appears to be inextricably linked to an interconnected network of conserved titratable residues that jointly contribute to the pH-dependent formation of the fusogenic spike. The cluster effect might also explain incongruous results in loss-of-function experiments which have shown that E proteins with an H144N mutation can still undergo the fusion event (*Nelson et al., 2009*), suggesting that His144 plays a part, but is not indispensable for triggering trimer formation.

Limitations

In this work, we modelled the portion of the flaviviral envelope with the aim of detecting cryptic pockets on its protein components, but we omitted 1–2 glycan molecules (at position N67 for DENV1-4 only, and at N153 for all addressed flaviviral species). At the time of conducting the research, the benzene mapping method had not been verified against systems containing sugar moieties, and we hypothesised that the shielding effect of glycans might hinder effective benzene exploration of the protein surface. Furthermore, the experimental data available regarding glycan types on each glycosylation site is scarce and is dependent on the viral species and the examined host (*Hacker et al., 2009*). However, the α pocket that contains the cluster of charges implicated in pH-dependent conformational change is located underneath the N153-glycan loop and as such, it is expected to at least partially affect the dynamics of pocket opening. This aspect of flaviviral envelope dynamics remains unaddressed at this time, and a topic for the future.

The membrane was modelled under geometric and compositional constraints. The curvature of the membrane was induced by introducing position restraints on $C\alpha$ atoms of the TM helices. The membrane consisted of three phospholipid species only, and it did not include sphingolipids or cholesterol, both of which are suggested to be functionally important for flaviviral infection (*Martín-Acebes et al., 2014; Carro and Damonte, 2013*). Taking into consideration the simplicity of the membrane composition model, as well as the application of position restraints in the lipid-proximal areas of proteins, detailed analysis of protein-lipid interactions was not pursued. Instead, the membrane served as a low-dielectric and solvent-excluding ‘shield’ that mimicked the approximate environment of the viral envelope and disallowed benzene to interact with the underside of the modelled raft.

The constant-pH simulations were performed on a truncated model of the E protein dimer, and the titrations were performed only on a small number of ionisable residues (specifically, the identified cluster charge residues and conserved histidines). The constant-pH method utilised the Generalised Born solvation model to attempt protonation, while the conformational sampling was performed in

an explicit water environment. These aspects of the constant-pH simulation setup may be limiting in terms of accuracy of the used model and the comprehensiveness of the protonation sampling. Therefore, the future continuation of this research may involve constant-pH simulation approaches that allow for simulations of bigger systems and with titration coordinates applied to all acid-ionisable residues, such as a recently developed method reported by [Aho et al., 2022](#).

Perspectives on cryptic hotspot discovery in flaviviruses

Prophylactic or therapeutic approaches for fighting flaviviral disease that utilise vaccines or antibodies might be suboptimal because of the danger of ADE. Instead, antiviral drugs that could target the flaviviral life cycle could be an effective alternative. Specifically, a potential drug targeting the early stages of the life cycle would not have to be membrane-soluble as it would target the viral particle in the extracellular environment. Furthermore, inhibiting the fusion step would contribute to prevention of ADE, because it would pre-empt viral escape from the endosome prior to degradation. Hence, antiviral drugs would not only offer protection on their own, but could also act in concert with vaccines. The inhibitory action of antiviral drugs by binding to the β -OG pocket has already been demonstrated for a range of flaviviruses ([de Wispelaere et al., 2018](#); [Li et al., 2019](#)), likely by preventing a hinge-like movement of the DI-DII interface necessary for formation of the fusogenic trimer. Similarly, targeting drug-like molecules to the α pocket could modulate charge cluster behaviour, effectively disrupting the pH-dependent dissociation of DI-DIII that is an essential step in the endosomal conformational change. Consistent with the cluster disruption and domain dissociation observed under low pH conditions, the surface area of the α pocket also increases with the net charge of the cluster ([Figure 6—figure supplement 4](#)). A drug molecule binding inside the pocket and stabilising it – both under neutral or low pH conditions – could conceivably prevent domain dissociation and pH-induced conformational changes of the viral envelope.

His144 is the most accessible to benzene probes in all examined flaviviral serotypes and thus likely to be a primary interaction moiety for potential antiviral drugs ([Figure 6—figure supplement 5A](#)). Our simulations show that benzene interactions with His144 also lead to an increased disorder within the cluster ([Figure 6—figure supplement 5](#)), which is an effect that would need to be considered in the context of drug design. Nonetheless, the proximity of other cluster residues suggests that all of them have the potential to play a role in drug-binding and in turn be modulated in their response to a low pH environment.

Conclusions

Our simulation studies of DENV1-4, YFV and Zika rafts provide a detailed perspective of the dynamic behaviour of flaviviral envelopes in the presence of benzene. We elaborated on a wider context of shared hotspot characteristics and potentially druggable cryptic pockets that could be utilised in the inhibition of the early stages of the viral life cycle. We have demonstrated that conserved surfaces share not only residue identity, but also similar patterns of benzene binding and increased SASA, highlighting the importance of residue conservation in the context of detection of shared binding hotspots. We also provided a detailed view of the β -OG pocket behaviour in different flaviviral species that exhibited a high degree of serotype specificity when examined for pocket opening events.

Furthermore, the α pocket was mapped across flaviviral serotypes and revealed a relatively consistent profile in terms of conservation, size and opening kinetics. It contained a cluster of conserved ionisable residues, and its position at the domain interfaces prompted further exploration of its role in pH-dependent conformational changes using constant-pH simulations. The conserved cluster displayed increased disorder at lower pH values and interdomain dissociation (a prerequisite for the formation of the fusogenic trimer). Protonation-driven disruption of the DI-DIII interface suggests its critical role in the endosomal pH-dependent conformational change. We also hypothesise that the cluster as a whole – as opposed to the usually considered histidine residues only – plays a role in the conformational shift, corroborating previous experimental observations that histidines alone might not be the only pH-sensing residues responsible for the orchestration of the pH-dependent conformational switch. Collectively, our findings provide a link between the sequence-dependent dynamic behaviour of hotspots and the mechanistic overview of their functional roles, and offer a comprehensive foundation for rational drug design targeting the envelope of flaviviruses.

Methods

Benzene mapping of flaviviral rafts

Raft modelling

A raft encompasses a large surface area of the viral envelope and therefore exhibits a distinct curvature in its quaternary structure, which manifests itself in angled contacts between neighbouring E proteins (**Figure 1B**). In comparison, crystal structures of ectodomain dimers (**Modis et al., 2003; Lu et al., 2019; Dai et al., 2016**) are invariably assembled outside the confines of a viral envelope, resulting in a 'flat' conformation of the dimeric structure. Such crystal structures are thus unable to fully replicate native interchain contacts present in the spherical particle, highlighting the necessity of using cryo-electron microscopy (cryo-EM) structures to produce a reliable raft assembly. Currently, there are only two published cryo-EM structures of the mature DENV particle at atomic resolution, both of which are of the DENV2 serotype (**Zhang et al., 2013a; Renner et al., 2021**). We selected a raft segment from a viral envelope cryo-EM structure solved at 3.6 Å resolution (PDB code: 3J27) and used it as a model for the DENV2 raft. We then created the remaining DENV serotypes (DENV1, 3, and 4) by mutating individual residues in Modeller 9.19 (**Sali and Blundell, 1993**), based on the DENV2 raft template. Sequence differences between YFV and DENV were deemed too great to create a reliable model of YFV based on the DENV2 template alone (**Figure 3—figure supplement 1B**). Instead, we used an existing crystal structure of the YFV ectodomain (PDB code: 6IW4; **Lu et al., 2019**) and modelled the missing loops (residues 269–272) with Modeller 9.19. Next, we used homology modelling based on the DENV2 template to create the missing TM regions and M proteins. Finally, the chains were aligned with the cryo-EM-derived DENV raft to reproduce the quaternary arrangement and curvature of a spherical viral particle. The Zika raft structure was taken from the existing cryo-EM Zika virus structure (PDB code: 6CO8) solved at 3.1 Å resolution¹⁰.

Protein rafts were embedded within a membrane (see below) and equilibrated during a series of simulations with position restraints of gradually decreasing force constants applied to the stem and TM regions of the proteins. Position restraints on the TM regions were necessary to retain the curvature of the raft. All structures were initially equilibrated for 200 ns with position restraints on the backbone atoms of stem and TM regions (E protein residues: 396–495; M protein residues: 21–72; numbering scheme corresponds to DENV2) using a force constant of 1000 kJ mol⁻¹ nm⁻², followed by another 200 ns equilibration with weaker position restraints ($F_c = 50$ kJ mol⁻¹ nm⁻²) on the TM region C α atoms only (E protein residues: 450–495; M protein residues: 39–72; numbering scheme corresponds to DENV2).

Membrane modelling

We used CHARMM-GUI (**Sunhwan et al., 2008; Wu et al., 2015**) to embed the DENV2 raft in a membrane containing 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC), 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE), and 1-palmitoyl-2-oleoyl-phosphatidylserine (POPS) lipids in 6:3:1 ratio, which reflected a simplified membrane composition of flaviviral envelopes expressed in a C6/36 mosquito cell line (**Perera et al., 2012; Zhang et al., 2012**). The produced membrane was initially modelled as a flat sheet, which resulted in lateral raft dimers being incorrectly submerged in the bilayer. In order to induce membrane curvature, we simulated the raft-membrane system with position restraints applied on all backbone atoms of the raft. Within ~400 ns of the simulation, the membrane gradually curved and accommodated the shape of the raft (**Figure 1—figure supplement 1**), with the TM regions transversing the bilayer and the amphipathic stem helices localising between hydrophilic headgroup and hydrophobic tail regions of the membrane (**Figure 1B**). In total, the membrane was equilibrated for 785 ns of unrestrained simulation time.

Benzene mapping simulation setup

The strategy using benzene probes for mapping pockets in membrane-bound proteins was adapted from our previous work (**Zuzic et al., 2020**). A massless virtual site was placed at the centre of mass of a benzene molecule, which was used as a point of repulsion between hydrophobic benzene molecules so as to prevent aggregation. Additionally, the same site was used as a point of repulsion between benzene probes and membrane lipids in order to hinder benzene sequestration into the bilayer. Benzene probes were added to a simulation box at 0.6 M concentration (1148 benzene molecules per

system). We tested this concentration for signs of protein unfolding or probe aggregation and found no evidence of either process occurring in our simulation systems (**Figure 2B**). The raft-membrane system containing benzene was solvated in TIP3P water (**Jorgensen et al., 1983**) and neutralised with NaCl, to a final salt concentration of 0.15 M.

Benzene mapping simulation protocol

All simulations were performed with the CHARMM36m force field (**Huang and MacKerell, 2013**) and in Gromacs 2018.2 simulation software (**Abraham et al., 2015**). Initially, systems were minimized using the steepest descent algorithm until forces converged or reached a maximum value of 100 kJ mol⁻¹ nm⁻¹. Temperature was set to 310 K and regulated using a velocity-rescaling thermostat (**Bussi et al., 2007**) with a time constant of 0.1 ps and applied on two coupled groups, one containing proteins and lipids, and the other, solute components. The Parrinello-Rahman barostat (**Parrinello and Rahman, 1981**) with semi-isotropic coupling and with a time constant of 2 ps was used to maintain a pressure of 1 atm. The compressibility was set to 4.5×10⁻⁵ bar⁻¹ for both x/y and z directions. The leapfrog algorithm integration step was set to 2 fs, while the LINCS constraint algorithm (**Hess et al., 1997**) was applied to all bonds associated with hydrogen atoms in the system. Computation of long-range electrostatic interactions was performed using the particle mesh Ewald (PME) method (**Darden et al., 1993; Essmann et al., 1995**), with a 1.2 nm cutoff for real-space interactions and with the integration of a Coulomb potential-shift modifier. The cutoff value for Van der Waals interactions was set to 1.2 nm, with a force-switch modifier applied at 1.0 nm.

The system was run in the canonical (NVT) ensemble for 1 ns and with strong position restraints ($F_c = 1000$ kJ mol⁻¹ nm⁻²) applied in all three dimensions to heavy atoms of both the protein and membrane components of the system. A subsequent 5-ns NPT equilibration involved position restraints on protein heavy atoms only, while the membrane was allowed to relax unrestrained. Finally, a production run was performed for 300 ns, with weak position restraints ($F_c = 50$ kJ mol⁻¹ nm⁻²) applied to the TM region C α atoms of the proteins (E protein residues: 450–495; M protein residues: 39–72; in DENV2 numbering system) in order to retain the curvature in the membrane model. A summary of all simulations and the corresponding number of repeats is listed in **Table 1**. In total, 14.4 μ s of production runs were generated.

Constant-pH MD simulations on DENV2 soluble E protein dimer

For constant-pH MD simulations we used a soluble portion of the E protein (sE) in a dimeric form (residues 1–395). The structure of the sE was taken from the DENV2 cryo-EM structure (PDB: 3J27)

Table 1. Flaviviral envelope raft simulations in presence or absence of benzene in the simulation system.

Strain	Solvent*	Time / ns	Repeats
DENV1-PVP159	water	300	3
DENV1-PVP159	0.6 M benzene	300	5
DENV2-NGC	water	300	3
DENV2-NGC	0.6 M benzene	300	5
DENV3-H87	water	300	3
DENV3-H87	0.6 M benzene	300	5
DENV4-H241	water	300	3
DENV4-H241	0.6 M benzene	300	5
YFV-17D	water	300	3
YFV-17D	0.6 M benzene	300	5
ZIKA-H/PF/2013	water	300	3
ZIKA-H/PF/2013	0.6 M benzene	300	5

*Both types of solvent contained charge-neutralising 0.15 M NaCl.

(Zhang *et al.*, 2013a). As the E protein was truncated for the purposes of simulation, we capped the C-terminus with an N-methyl amide (NME) group. Residues His27, Asp42, His144, His244, His317, and Glu368 in both chains were selected to titrate. His282, although present in the structure and implicated in pH-dependent conformational change, was excluded from the selection as it is natively interacting with the stem region (not present in this setup) and was therefore in a biologically irrelevant environment. Native disulfide bonds were included in the structure. The sE dimer was placed in an octahedron box and solvated with ~84,000 TIP3P water molecules (Jorgensen *et al.*, 1983) and neutralising 0.15 M NaCl.

The simulations were performed with the Amber ff99SB force field (Hornak *et al.*, 2006) modified to include additional parameters relevant for constant-pH. Bond radii of Asp and Glu residues were reduced as carboxylate oxygens were defined with two hydrogens on each protonation site. All constant-pH simulations were performed using the Amber20 simulation software (Case *et al.*, 2005). The system was initially minimised using a combination of steepest descent and conjugate gradient algorithms performed in 5,000 cycles and with protein backbone position restraints defined with a force constant of 4184 kJ mol⁻¹nm⁻² (10 kcal mol⁻¹Å⁻²). The system was then gradually heated to 310 K over 400 ps by applying Langevin thermostat with a collision frequency of 5.0 ps⁻¹. The NPT equilibration step was performed for 4 ns with Langevin dynamics maintaining a constant pressure of 1 atm defined by a collision frequency of 1.0 ps⁻¹. The production steps of constant-pH simulations were run under 10 pH conditions across the scale of pH = 0, 1, 2, ... 9. Production runs were simulated for 22 ns each in the NPT ensemble, with attempted protonation state changes every 100 simulation steps (0.2 ps). Protonation attempts were performed by Monte Carlo sampling (Mongan *et al.*, 2004) based on transition free energies derived from generalised Born electrostatics (Bashford and Case, 2000) and with the Metropolis criterion used to determine if the protonation change should be accepted or not. All simulations were run with a 2 fs integration step, applying a SHAKE constraint algorithm (Ryckaert *et al.*, 1977) to all bonds involving hydrogen atoms. Short-range electrostatic and Van der Waals interaction cutoffs were both set to 0.8 nm. Long-range electrostatics interactions were calculated using the PME method (Darden *et al.*, 1993; Essmann *et al.*, 1995).

Constant-pH simulations were run in quadruplicates, resulting in a total production time of 880 ns.

Analysis procedures

Sequence analysis

Flaviviral polyprotein sequences were accessed from GenBank using following accession codes: DENV1 strain PVP159: AEM92304.1; DENV2 strain NGS-C: P14340.2; DENV3 strain H87: P27915.1; DENV4 strain H241: Q58HT7.1; YFV strain 17D: P03314.1; Zika strain H/PF/2013: A0A024B7W1.1. Sequence similarity was determined using the Sequence Identity and Similarity (SIAS) software. Multiple sequence alignment of the E and M proteins was performed using MUSCLE (Edgar, 2004). Alignment visualisation and generation of the neighbour-joining (NJ) tree was performed using Jalview 2.11.1.4 (Waterhouse *et al.*, 2009). Conservation assessment and scoring was carried out using the ConSurf Server (Ashkenazy *et al.*, 2016), where a multiple sequence alignment was generated using the HMMER homolog search algorithm in the UNIREF-90 database with an E-value cutoff of 0.0001.

MD simulations analysis

After removing the position restraints from benzene probes, the system required additional equilibration time during which benzene established interactions with protein surfaces. We therefore excluded the first 50 ns of all production runs from the analysis. Simulation properties, including solvent-accessible surface area (SASA) (Eisenhaber *et al.*, 1995), secondary structure content (Kabsch and Sander, 1983), radius of gyration, select residue distances, root-mean-square deviations (RMSDs) and root-mean-square fluctuations (RMSFs), were analysed using Gromacs 2018.2 analysis tools (Abraham *et al.*, 2015). Contact distances between benzene and residues of interest were calculated using VMD 1.9.3 (Humphrey *et al.*, 1996). In order to avoid counting contacts with the β-OG pocket residues that occur when the probes are outside the binding site, we counted only benzene proximal to at least two oppositely positioned pocket residues. The initial detection of protein pockets was performed with the MDpocket software (Schmidtke *et al.*, 2011). Correlation of raft pocket densities across serotypes was calculated using the Fit in Map tool in ChimeraX 1.0 (Pettersen *et al.*, 2021).

Protonation states of all titrated residues were extracted using the *chpstats* tool implemented in AmberTools21 (Case *et al.*, 2005). pK_a predictions of titratable residues based on a static structure were calculated on the DENV2 E proteins (PDB: 3J27) (Zhang *et al.*, 2013a) in an asymmetric unit using the Finite difference Poisson-Boltzmann/Debye-Hückel (FD/DH) method (Warwicker, 2004).

Statistical analysis

All residue-based analyses across serotypes were performed after alignment, ensuring that the residues corresponded to their shared evolutionary origin. All correlation analyses were calculated with a non-parametric Spearman's rank correlation method. β -OG pocket clustering was performed with a *k*-means clustering algorithm ($k=2$) based on the measurements of pocket SASA and the number of contacts established between all the residues lining the pocket. All pairwise significance levels were determined with a Wilcoxon signed-rank test used in comparison of two non-parametric samples. Cluster analysis of the charge cluster residues, applied to all atoms within the group, was based on the *gromos* algorithm (Daura *et al.*, 1999) implemented in Gromacs with a specified cutoff value of 0.26 nm. Principal component analysis (PCA) of the charge cluster descriptors – all combinations of intra-cluster distances – was performed on a scaled and centered sample containing all structures across pH values, chains, and constant-pH simulation repeats. Visualisation of the original variable axes used the *ggbiplot* package in R 3.6.3. Additional data processing, analysis and visualisation were carried out in R 3.6.3.

Acknowledgements

The research was supported by the A*STAR Bioinformatics Institute (BII) core funds, the A*STAR Graduate Academy (A*GA), and the University of Manchester. PJB, JKM, and GSA acknowledge NRF (NRF2017NRF-CRP001-027) for funding. The computational work for this article was performed on resources of the National Supercomputing Centre, Singapore (<https://www.nsc.sg>), A*STAR Computational Resource Centre, A*STAR Bioinformatics Institute, and the N8 Centre of Excellence in Computationally Intensive Research (N8 CIR) provided and funded by the N8 research partnership and EPSRC (Grant No. EP/T022167/1). We are grateful to Ana Damjanovic for helpful advice regarding constant-pH simulations.

Additional information

Funding

Funder	Grant reference number	Author
National Research Foundation Singapore	NRF2017NRF-CRP001-027	Jan K Marzinek Ganesh S Anand Peter J Bond
Agency for Science, Technology and Research		Lorena Zuzic Jan K Marzinek Peter J Bond
University of Manchester		Lorena Zuzic Jim Warwicker

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Lorena Zuzic, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing – review and editing; Jan K Marzinek, Conceptualization, Supervision, Validation, Investigation, Visualization, Writing – review and editing; Ganesh S Anand, Conceptualization, Writing – review and editing; Jim Warwicker, Conceptualization, Supervision, Funding acquisition, Writing – review and editing; Peter J Bond, Conceptualization, Resources, Data curation, Supervision, Validation, Visualization, Writing - original draft, Project administration, Writing – review and editing

Author ORCIDsLorena Zuzic  <http://orcid.org/0000-0002-7834-612X>Jan K Marzinek  <http://orcid.org/0000-0002-5493-8753>Ganesh S Anand  <http://orcid.org/0000-0001-8995-3067>Peter J Bond  <http://orcid.org/0000-0003-2900-098X>**Decision letter and Author response**Decision letter <https://doi.org/10.7554/eLife.82447.sa1>Author response <https://doi.org/10.7554/eLife.82447.sa2>

Additional files**Supplementary files**

- MDAR checklist

Data availability

The full simulation trajectories have not been uploaded to a sharing server due to the size of the datasets (several terabytes of data in total). However, we have uploaded all simulation setup data such that a researcher could re-run our calculations (i.e. coordinates including initial and final frames, parameters, checkpoints, and portable binary run input files, for a total of 88 simulation systems) at Zenodo: <https://doi.org/10.5281/zenodo.7037906>. More extensive simulation trajectory files will be made available upon request to the corresponding author, without the need to e.g. submit a project proposal. There will be no restrictions on who can access this data. Code/software used to analyse the data are freely available and in the public domain.

References

- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. 2015. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**:19–25. DOI: <https://doi.org/10.1016/j.softx.2015.06.001>
- Acosta EG, Castilla V, Damonte EB. 2008. Functional entry of dengue virus into *Aedes albopictus* mosquito cells is dependent on clathrin-mediated endocytosis. *The Journal of General Virology* **89**:474–484. DOI: <https://doi.org/10.1099/vir.0.83357-0>, PMID: 18198378
- Aho N, Buslaev P, Jansen A, Bauer P, Groenhof G, Hess B. 2022. Scalable constant pH molecular dynamics in GROMACS. *Journal of Chemical Theory and Computation* **18**:6148–6160. DOI: <https://doi.org/10.1021/acs.jctc.2c00516>
- Allison SL, Schalich J, Stiasny K, Mandl CW, Heinz FX. 2001. Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. *Journal of Virology* **75**:4268–4275. DOI: <https://doi.org/10.1128/JVI.75.9.4268-4275.2001>, PMID: 11287576
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **44**:W344–W350. DOI: <https://doi.org/10.1093/nar/gkw408>, PMID: 27166375
- Balsitis SJ, Williams KL, Lachica R, Flores D, Kyle JL, Mehlhop E, Johnson S, Diamond MS, Beatty PR, Harris E. 2010. Lethal antibody enhancement of dengue disease in mice is prevented by Fc modification. *PLoS Pathogens* **6**:e1000790. DOI: <https://doi.org/10.1371/journal.ppat.1000790>, PMID: 20168989
- Barba-Spaeth G, Dejnirattisai W, Rouvinski A, Vaney MC, Medits I, Sharma A, Simon-Lorière E, Sakuntabhai A, Cao-Lormeau VM, Haouz A, England P, Stiasny K, Mongkolsapaya J, Heinz FX, Screaton GR, Rey FA. 2016. Structural basis of potent zika–dengue virus antibody cross-neutralization. *Nature* **536**:48–53. DOI: <https://doi.org/10.1038/nature18938>
- Bashford D, Case DA. 2000. Generalized born models of macromolecular solvation effects. *Annual Review of Physical Chemistry* **51**:129–152. DOI: <https://doi.org/10.1146/annurev.physchem.51.1.129>, PMID: 11031278
- Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GRW, Simmons CP, Scott TW, Farrar JJ, Hay SI. 2013. The global distribution and burden of dengue. *Nature* **496**:504–507. DOI: <https://doi.org/10.1038/nature12060>, PMID: 23563266
- Boldescu V, Behnam MAM, Vasilakis N, Klein CD. 2017. Broad-Spectrum agents for flaviviral infections: dengue, Zika and beyond. *Nature Reviews. Drug Discovery* **16**:565–586. DOI: <https://doi.org/10.1038/nrd.2017.33>, PMID: 28473729
- Brady OJ, Gething PW, Bhatt S, Messina JP, Brownstein JS, Hoen AG, Moyes CL, Farlow AW, Scott TW, Hay SI. 2012. Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Neglected Tropical Diseases* **6**:e1760. DOI: <https://doi.org/10.1371/journal.pntd.0001760>, PMID: 22880140
- Bussi G, Donadio D, Parrinello M. 2007. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126**:014101. DOI: <https://doi.org/10.1063/1.2408420>

- Carro AC**, Damonte EB. 2013. Requirement of cholesterol in the viral envelope for dengue virus infection. *Virus Research* **174**:78–87. DOI: <https://doi.org/10.1016/j.virusres.2013.03.005>, PMID: 23517753
- Case DA**, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**:1668–1688. DOI: <https://doi.org/10.1002/jcc.20290>, PMID: 16200636
- Cheng AC**, Coleman RG, Smyth KT, Cao Q, Souillard P, Caffrey DR, Salzberg AC, Huang ES. 2007. Structure-Based maximal affinity model predicts small-molecule druggability. *Nature Biotechnology* **25**:71–75. DOI: <https://doi.org/10.1038/nbt1273>, PMID: 17211405
- Dai L**, Song J, Lu X, Deng YQ, Musyoki AM, Cheng H, Zhang Y, Yuan Y, Song H, Haywood J, Xiao H, Yan J, Shi Y, Qin CF, Qi J, Gao GF. 2016. Structures of the Zika virus envelope protein and its complex with a flavivirus broadly protective antibody. *Cell Host & Microbe* **19**:696–704. DOI: <https://doi.org/10.1016/j.chom.2016.04.013>
- Darden T**, York D, Pedersen L. 1993. Particle mesh ewald: an N-log (N) method for ewald sums in large systems. *The Journal of Chemical Physics* **98**:12. DOI: <https://doi.org/10.1063/1.464397>
- Daura X**, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. 1999. Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition* **38**:236–240. DOI: [https://doi.org/10.1002/\(SICI\)1521-3773\(19990115\)38:1/2<236::AID-ANIE236>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M)
- DeLano WL**. 2002. Unraveling hot spots in binding interfaces: progress and challenges. *Current Opinion in Structural Biology* **12**:14–20. DOI: [https://doi.org/10.1016/S0959-440x\(02\)00283-x](https://doi.org/10.1016/S0959-440x(02)00283-x), PMID: 11839484
- de Wispelaere M**, Lian W, Potisopon S, Li P-C, Jang J, Ficarro SB, Clark MJ, Zhu X, Kaplan JB, Pitts JD, Wales TE, Wang J, Engen JR, Marto JA, Gray NS, Yang PL. 2018. Inhibition of flaviviruses by targeting a conserved pocket on the viral envelope protein. *Cell Chemical Biology* **25**:1006–1016. DOI: <https://doi.org/10.1016/j.chembiol.2018.05.011>, PMID: 29937406
- Edgar RC**. 2004. Muscle: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**:113. DOI: <https://doi.org/10.1186/1471-2105-5-113>, PMID: 15318951
- Eisenhaber F**, Lijnzaad P, Argos P, Sander C, Scharf M. 1995. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of Computational Chemistry* **16**:273–284. DOI: <https://doi.org/10.1002/jcc.540160303>
- Essmann U**, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh ewald method. *The Journal of Chemical Physics* **103**:8577–8593. DOI: <https://doi.org/10.1063/1.470117>
- Fibriansah G**, Ng TS, Kostyuchenko VA, Lee J, Lee S, Wang J, Lok SM. 2013. Structural changes in Dengue virus when exposed to a temperature of 37 C. *Journal of Virology* **87**:7585–7592. DOI: <https://doi.org/10.1128/JVI.00757-13>, PMID: 23637405
- Fibriansah G**, Lim EXY, Marzinek JK, Ng TS, Tan JL, Huber RG, Lim XN, Chew VSY, Kostyuchenko VA, Shi J, Anand GS, Bond PJ, Crowe JE, Lok SM. 2021. Antibody affinity versus dengue morphology influences neutralization. *PLOS Pathogens* **17**:e1009331. DOI: <https://doi.org/10.1371/journal.ppat.1009331>, PMID: 33621239
- Fritz R**, Stiasny K, Heinz FX. 2008. Identification of specific histidines as pH sensors in flavivirus membrane fusion. *The Journal of Cell Biology* **183**:353–361. DOI: <https://doi.org/10.1083/jcb.200806081>, PMID: 18936253
- García-Moreno BE**, Dwyer JJ, Gittis AG, Lattman EE, Spencer DS, Stites WE. 1997. Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophysical Chemistry* **64**:211–224. DOI: [https://doi.org/10.1016/S0301-4622\(96\)02238-7](https://doi.org/10.1016/S0301-4622(96)02238-7)
- Gaunt MW**, Sall AA, de Lamballerie X, Falconar AKI, Dzhanian TI, Gould EA. 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology. *The Journal of General Virology* **82**:88. DOI: <https://doi.org/10.1099/0022-1317-82-8-1867>
- Gould EA**. 1998. Dengue haemorrhagic fever: diagnosis, treatment, prevention and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**:470. DOI: [https://doi.org/10.1016/S0035-9203\(98\)91100-2](https://doi.org/10.1016/S0035-9203(98)91100-2)
- Guvench O**, MacKerell AD. 2009. Computational fragment-based binding site identification by ligand competitive saturation. *PLOS Computational Biology* **5**:e1000435. DOI: <https://doi.org/10.1371/journal.pcbi.1000435>, PMID: 19593374
- Guy B**, Noriega F, Ochiai RL, L'azou M, Delore V, Skipetrova A, Verdier F, Coudeville L, Savarino S, Jackson N. 2017. A recombinant live attenuated tetravalent vaccine for the prevention of dengue. *Expert Review of Vaccines* **16**:1–13. DOI: <https://doi.org/10.1080/14760584.2017.1335201>, PMID: 28590795
- Guzman MG**, Alvarez M, Halstead SB. 2013. Secondary infection as a risk factor for dengue hemorrhagic fever/dengue shock syndrome: an historical perspective and role of antibody-dependent enhancement of infection. *Archives of Virology* **158**:1445–1459. DOI: <https://doi.org/10.1007/s00705-013-1645-3>
- Hacker K**, White L, de Silva AM. 2009. N-Linked glycans on dengue viruses grown in mammalian and insect cells. *Journal of General Virology* **90**:2097–2106. DOI: <https://doi.org/10.1099/vir.0.012120-0>
- Halstead SB**. 1979. In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody. *Journal of Infectious Diseases* **140**:527–533. DOI: <https://doi.org/10.1093/infdis/140.4.527>
- Harrison SC**. 2015. Viral membrane fusion. *Virology* **479–480**:498–507. DOI: <https://doi.org/10.1016/j.virol.2015.03.043>, PMID: 25866377
- Heinz FX**, Stiasny K, Püschner-Auer G, Holzmann H, Allison SL, Mandl CW, Kunz C. 1994. Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the Heterodimeric association with protein Prm. *Virology* **198**:109–117. DOI: <https://doi.org/10.1006/viro.1994.1013>, PMID: 8259646

- Hess B, Bekker H, Berendsen HJC, Fraaije J. 1997. LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**:1463–1472. DOI: [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
- Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. 2006. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**:712–725. DOI: <https://doi.org/10.1002/prot.21123>
- Huang J, MacKerell AD. 2013. Charmm36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry* **34**:2135–2145. DOI: <https://doi.org/10.1002/jcc.23354>, PMID: 23832629
- Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**:33–38. DOI: [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5), PMID: 8744570
- Isom DG, Castañeda CA, Cannon BR, García-Moreno E. B. 2011. Large shifts in pK a values of lysine residues buried inside a protein. *PNAS* **108**:5260–5265. DOI: <https://doi.org/10.1073/pnas.1010750108>
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**:926–935. DOI: <https://doi.org/10.1063/1.445869>
- Junjhon J, Edwards TJ, Utaipat U, Bowman VD, Holdaway HA, Zhang W, Keelapang P, Puttikhunt C, Perera R, Chipman PR, Kasinrerak W, Malasit P, Kuhn RJ, Sittisombut N. 2010. Influence of PR-M cleavage on the heterogeneity of extracellular Dengue virus particles. *Journal of Virology* **84**:8353–8358. DOI: <https://doi.org/10.1128/JVI.00696-10>, PMID: 20519400
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**:2577–2637. DOI: <https://doi.org/10.1002/bip.360221211>, PMID: 6667333
- Kamal M, Kenawy MA, Rady MH, Khaled AS, Samy AM. 2018. Mapping the global potential distributions of two arboviral vectors *Aedes aegypti* and *Ae. albopictus* under changing climate. *PLOS ONE* **13**:e0210122. DOI: <https://doi.org/10.1371/journal.pone.0210122>, PMID: 30596764
- Kanai R, Kar K, Anthony K, Gould LH, Ledizet M, Fikrig E, Marasco WA, Koski RA, Modis Y. 2006. Crystal structure of West Nile virus envelope glycoprotein reveals viral surface epitopes. *Journal of Virology* **80**:11000–11008. DOI: <https://doi.org/10.1128/JVI.01735-06>, PMID: 16943291
- Kolb P, Cafilisch A. 2006. Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *Journal of Medicinal Chemistry* **49**:7384–7392. DOI: <https://doi.org/10.1021/jm060838i>
- Kostyuchenko VA, Zhang Q, Tan JL, Ng TS, Lok SM. 2013. Immature and mature dengue serotype 1 virus structures provide insight into the maturation process. *Journal of Virology* **87**:7700–7707. DOI: <https://doi.org/10.1128/JVI.00197-13>, PMID: 23637416
- Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Van Bortel W, Hendrickx G, Schaffner F, Elyazar IRF, Teng H-J, Brady OJ, Messina JP, Pigott DM, Scott TW, Smith DL, Wint GRW, et al. 2015. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife* **4**:e08347. DOI: <https://doi.org/10.7554/eLife.08347>, PMID: 26126267
- Kraemer MUG, Reiner RC, Brady OJ, Messina JP, Gilbert M, Pigott DM, Yi D, Johnson K, Earl L, Marczak LB, Shirude S, Davis Weaver N, Bisanzio D, Perkins TA, Lai S, Lu X, Jones P, Coelho GE, Carvalho RG, Van Bortel W, et al. 2019. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology* **4**:854–863. DOI: <https://doi.org/10.1038/s41564-019-0376-y>
- Kuzmanic A, Bowman GR, Juarez-Jimenez J, Michel J, Gervasio FL. 2020. Investigating cryptic binding sites by molecular dynamics simulations. *Accounts of Chemical Research* **53**:654–661. DOI: <https://doi.org/10.1021/acs.accounts.9b00613>
- Li P-C, Jang J, Hsia C-Y, Grooms PV, Lian W, de Wispelaere M, Pitts JD, Wang J, Kwiatkowski N, Gray NS, Yang PL. 2019. Small molecules targeting the flavivirus E protein with broad-spectrum activity and antiviral efficacy in vivo. *ACS Infectious Diseases* **5**:460–472. DOI: <https://doi.org/10.1021/acsinfecdis.8b00322>
- Lim XX, Chandramohan A, Lim XYE, Bag N, Sharma KK, Wirawan M, Wohland T, Lok SM, Anand GS. 2017. Conformational changes in intact Dengue virus reveal Serotype-specific expansion. *Nature Communications* **8**:14339. DOI: <https://doi.org/10.1038/ncomms14339>, PMID: 28186093
- Lim X-N, Shan C, Marzinek JK, Dong H, Ng TS, Ooi JSG, Fibriansah G, Wang J, Verma CS, Bond PJ, Shi P-Y, Lok S-M. 2019. Molecular basis of dengue virus serotype 2 morphological switch from 29°C to 37°C. *PLOS Pathogens* **15**:e1007996. DOI: <https://doi.org/10.1371/journal.ppat.1007996>, PMID: 31536610
- Lu X, Xiao H, Li S, Qi J, Shi Y, Gao GF, Lu X, Xiao H, Li S, Pang X, Song J, Liu S, Cheng H, Li Y, Wang X, Huang C, Guo T, ter Meulen J, Daffis S, Yan J, et al. 2019. Double lock of a human neutralizing and protective monoclonal antibody targeting the yellow fever virus envelope. *Cell Reports* **26**:438–446. DOI: <https://doi.org/10.1016/j.celrep.2018.12.065>, PMID: 30625326
- Luca VC, AbiMansour J, Nelson CA, Fremont DH. 2012. Crystal structure of the Japanese encephalitis virus envelope protein. *Journal of Virology* **86**:2337–2346. DOI: <https://doi.org/10.1128/JVI.06072-11>
- Martín-Acebes MA, Merino-Ramos T, Blázquez A-B, Casas J, Escribano-Romero E, Sobrino F, Saiz J-C, Dermody TS. 2014. The composition of West Nile virus lipid envelope unveils a role of sphingolipid metabolism in flavivirus biogenesis. *Journal of Virology* **88**:12041–12054. DOI: <https://doi.org/10.1128/JVI.02061-14>
- Mattos C, Ringe D. 1996. Locating and characterizing binding sites on proteins. *Nature Biotechnology* **14**:595–599. DOI: <https://doi.org/10.1038/nbt0596-595>, PMID: 9630949

- Modis Y**, Ogata S, Clements D, Harrison SC. 2003. A ligand-binding pocket in the Dengue virus envelope glycoprotein. *PNAS* **100**:6986–6991. DOI: <https://doi.org/10.1073/pnas.0832193100>, PMID: 12759475
- Modis Y**, Ogata S, Clements D, Harrison SC. 2004. Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**:313–319. DOI: <https://doi.org/10.1038/nature02165>, PMID: 14737159
- Modis Y**, Ogata S, Clements D, Harrison SC. 2005. Variable surface Epitopes in the crystal structure of Dengue virus type 3 envelope glycoprotein. *Journal of Virology* **79**:1223–1231. DOI: <https://doi.org/10.1128/JVI.79.2.1223>
- Mongan J**, Case DA, McCammon JA. 2004. Constant pH molecular dynamics in generalized born implicit solvent. *Journal of Computational Chemistry* **25**:2038–2048. DOI: <https://doi.org/10.1002/jcc.20139>, PMID: 15481090
- Morrone SR**, Lok SM. 2019. Structural perspectives of antibody-dependent enhancement of infection of dengue virus. *Current Opinion in Virology* **36**:1–8. DOI: <https://doi.org/10.1016/j.coviro.2019.02.002>, PMID: 30844538
- Morrone SR**, Chew VSY, Lim XN, Ng TS, Kostyuchenko VA, Zhang S, Wirawan M, Chew PL, Lee J, Tan JL, Wang J, Tan TY, Shi J, Screaton G, Morais MC, Lok SM. 2020. High flavivirus structural plasticity demonstrated by a non-spherical morphological variant. *Nature Communications* **11**:3112. DOI: <https://doi.org/10.1038/s41467-020-16925-y>, PMID: 32561757
- Mosso C**, Galván-Mendoza IJ, Ludert JE, del Angel RM. 2008. Endocytic pathway followed by dengue virus to infect the mosquito cell line C6/36 HT. *Virology* **378**:193–199. DOI: <https://doi.org/10.1016/j.virol.2008.05.012>, PMID: 18571214
- Nelson S**, Poddar S, Lin TY, Pierson TC. 2009. Protonation of individual histidine residues is not required for the pH-dependent entry of West Nile virus: evaluation of the "histidine switch" hypothesis. *Journal of Virology* **83**:12631–12635. DOI: <https://doi.org/10.1128/JVI.01072-09>, PMID: 19776132
- Nybakken GE**, Nelson CA, Chen BR, Diamond MS, Fremont DH. 2006. Crystal structure of the West Nile virus envelope glycoprotein. *Journal of Virology* **80**:11467–11474. DOI: <https://doi.org/10.1128/JVI.01125-06>, PMID: 16987985
- Oleinikovas V**, Saladino G, Cossins BP, Gervasio FL. 2016. Understanding cryptic pocket formation in protein targets by enhanced sampling simulations. *Journal of the American Chemical Society* **138**:14257–14263. DOI: <https://doi.org/10.1021/jacs.6b05425>
- Parrinello M**, Rahman A. 1981. Polymorphic transitions in single crystals: a new molecular dynamics method. *Journal of Applied Physics* **52**:7182–7190. DOI: <https://doi.org/10.1063/1.328693>
- Perera R**, Riley C, Isaac G, Hopf-Jannasch AS, Moore RJ, Weitz KW, Pasa-Tolic L, Metz TO, Adamec J, Kuhn RJ. 2012. Dengue virus infection perturbs lipid homeostasis in infected mosquito cells. *PLOS Pathogens* **8**:e1002584. DOI: <https://doi.org/10.1371/journal.ppat.1002584>, PMID: 22457619
- Pettersen EF**, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF chimeraX: structure visualization for researchers, educators, and developers. *Protein Science* **30**:70–82. DOI: <https://doi.org/10.1002/pro.3943>, PMID: 32881101
- Randolph VB**, Stollar V. 1990. Low PH-induced cell fusion in Flavivirus-infected Aedes Albopictus cell cultures. *The Journal of General Virology* **71** (Pt 8):1845–1850. DOI: <https://doi.org/10.1099/0022-1317-71-8-1845>, PMID: 2167941
- Renner M**, Dejnirattisai W, Carrique L, Martin IS, Karia D, Ilca SL, Ho SF, Kotecha A, Keown JR, Mongkolsapaya J, Screaton GR, Grimes JM. 2021. Flavivirus maturation leads to the formation of an occupied lipid pocket in the surface Glycoproteins. *Nature Communications* **12**:1238. DOI: <https://doi.org/10.1038/s41467-021-21505-9>, PMID: 33623019
- Rey FA**, Stiasny K, Vaney M-C, Dellarole M, Heinz FX. 2018. The bright and the dark side of human antibody responses to flaviviruses: lessons for vaccine design. *EMBO Reports* **19**:206–224. DOI: <https://doi.org/10.15252/embr.201745302>, PMID: 29282215
- Ryckaert JP**, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**:327–341. DOI: [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5)
- Sali A**, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**:779–815. DOI: <https://doi.org/10.1006/jmbi.1993.1626>, PMID: 8254673
- Sazanavets I**, Warwicker J. 2015. Computational tools for interpreting ion channel PH-dependence. *PLOS ONE* **10**:e0125293. DOI: <https://doi.org/10.1371/journal.pone.0125293>, PMID: 25915903
- Schmidt D**, Boehm M, McClendon CL, Torella R, Gohlke H. 2019. Cosolvent-enhanced sampling and Unbiased identification of cryptic pockets suitable for structure-based drug design. *Journal of Chemical Theory and Computation* **15**:3331–3343. DOI: <https://doi.org/10.1021/acs.jctc.8b01295>, PMID: 30998331
- Schmidtke P**, Barril X. 2010. A high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry* **53**:574. DOI: <https://doi.org/10.1021/jm100574m>
- Schmidtke P**, Bidon-Chanal A, Luque FJ, Barril X. 2011. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **27**:3276–3285. DOI: <https://doi.org/10.1093/bioinformatics/btr550>, PMID: 21967761
- Sevvana M**, Long F, Miller AS, Klose T, Buda G, Sun L, Kuhn RJ, Rossmann MG. 2018. Refinement and analysis of the mature Zika virus Cryo-EM structure at 3.1 Å resolution. *Structure* **26**:1169–1177. DOI: <https://doi.org/10.1016/j.str.2018.05.006>, PMID: 29958768
- Sirohi D**, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, Kuhn RJ. 2016. The 3.8 Å resolution cryo-EM structure of Zika virus. *Science* **352**:467–470. DOI: <https://doi.org/10.1126/science.aaf5316>, PMID: 27033547

- Srivastava J**, Barber DL, Jacobson MP. 2007. Intracellular PH sensors: Design principles and functional significance. *Physiology* **22**:30–39. DOI: <https://doi.org/10.1152/physiol.00035.2006>, PMID: 17289928
- Stiasny K**, Fritz R, Pangerl K, Heinz FX. 2011. Molecular mechanisms of flavivirus membrane fusion. *Amino Acids* **41**:1159–1163. DOI: <https://doi.org/10.1007/s00726-009-0370-4>, PMID: 19882217
- Sunhwan J**, Taehoon K, Vidyashankara IG, Wonpil I. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **29**:1859–1865. DOI: <https://doi.org/10.1002/jcc>
- Tan YS**, Reeks J, Brown CJ, Thean D, Ferrer Gago FJ, Yuen TY, Goh ETL, Lee XEC, Jennings CE, Joseph TL, Lakshminarayanan R, Lane DP, Noble MEM, Verma CS. 2016. Benzene probes in molecular Dynamics Simulations reveal novel binding sites for ligand design. *The Journal of Physical Chemistry Letters* **7**:3452–3457. DOI: <https://doi.org/10.1021/acs.jpcllett.6b01525>, PMID: 27532490
- Warwicker J**. 2004. Improved PKA calculations through flexibility based sampling of a water-dominated interaction scheme. *Protein Science* **13**:2793–2805. DOI: <https://doi.org/10.1110/ps.04785604>, PMID: 15388865
- Waterhouse AM**, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview version 2 -- a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**:1189–1191. DOI: <https://doi.org/10.1093/bioinformatics/btp033>, PMID: 19151095
- Williams KL**, Sukupolvi-Petty S, Beltramello M, Johnson S, Sallusto F, Lanzavecchia A, Diamond MS, Harris E. 2013. Therapeutic efficacy of antibodies lacking Fcγ receptor binding against lethal dengue virus infection is due to neutralizing potency and blocking of enhancing antibodies [corrected]. *PLoS Pathogens* **9**:e1003157. DOI: <https://doi.org/10.1371/journal.ppat.1003157>, PMID: 23459315
- Wirawan M**, Fibriansah G, Marzinek JK, Lim XX, Ng TS, Sim AYL, Zhang Q, Kostyuchenko VA, Shi J, Smith SA, Verma CS, Anand G, Crowe JE, Bond PJ, Lok SM. 2019. Mechanism of enhanced immature dengue virus attachment to endosomal membrane induced by PRM antibody. *Structure* **27**:253–267. DOI: <https://doi.org/10.1016/j.str.2018.10.009>, PMID: 30471923
- Wu EL**, Cheng X, Jo S, Rui H, Song KC, Dávila-Contreras EM, Qi Y, Lee J, Monje-galvan V, Venable RM, Jeffery KB, Wonpil I. 2015. CHARMM-GUI membrane builder toward realistic biological membrane Simulations. *Journal of Computational Chemistry* **35**:1997–2004. DOI: <https://doi.org/10.1002/jcc.23702>. CHARMM-GUI
- Zerbe BS**, Hall DR, Vajda S, Whitty A, Kozakov D. 2012. Relationship between hot spot residues and ligand binding hot spots in protein–protein interfaces. *Journal of Chemical Information and Modeling* **52**:2236–2244. DOI: <https://doi.org/10.1021/ci300175u>
- Zhang Y**, Zhang W, Ogata S, Clements D, Strauss JH, Baker TS, Kuhn RJ, Rossmann MG. 2004. Conformational changes of the flavivirus E glycoprotein. *Structure* **12**:1607–1618. DOI: <https://doi.org/10.1016/j.str.2004.06.019>
- Zhang Q**, Hunke C, Yau YH, Seow V, Lee S, Tanner LB, Guan XL, Wenk MR, Fibriansah G, Chew PL, Kukkaro P, Biuković G, Shi P-Y, Shochat SG, Grüber G, Lok S-M. 2012. The stem region of premembrane protein plays an important role in the virus surface protein rearrangement during dengue maturation. *Journal of Biological Chemistry* **287**:40525–40534. DOI: <https://doi.org/10.1074/jbc.M112.384446>
- Zhang X**, Ge P, Yu X, Brannan JM, Bi G, Zhang Q, Schein S, Zhou ZH. 2013a. Cryo-Em structure of the mature dengue virus at 3.5-Å resolution. *Nature Structural & Molecular Biology* **20**:105–110. DOI: <https://doi.org/10.1038/nsmb.2463>
- Zhang X**, Sheng J, Plevka P, Kuhn RJ, Diamond MS, Rossmann MG. 2013b. Dengue structure differs at the temperatures of its human and mosquito hosts. *PNAS* **110**:6795–6799. DOI: <https://doi.org/10.1073/pnas.1304300110>
- Zuzic L**, Marzinek JK, Warwicker J, Bond PJ. 2020. A benzene-mapping approach for Uncovering cryptic pockets in membrane-bound proteins. *Journal of Chemical Theory and Computation* **16**:5948–5959. DOI: <https://doi.org/10.1021/acs.jctc.0c00370>, PMID: 32786908